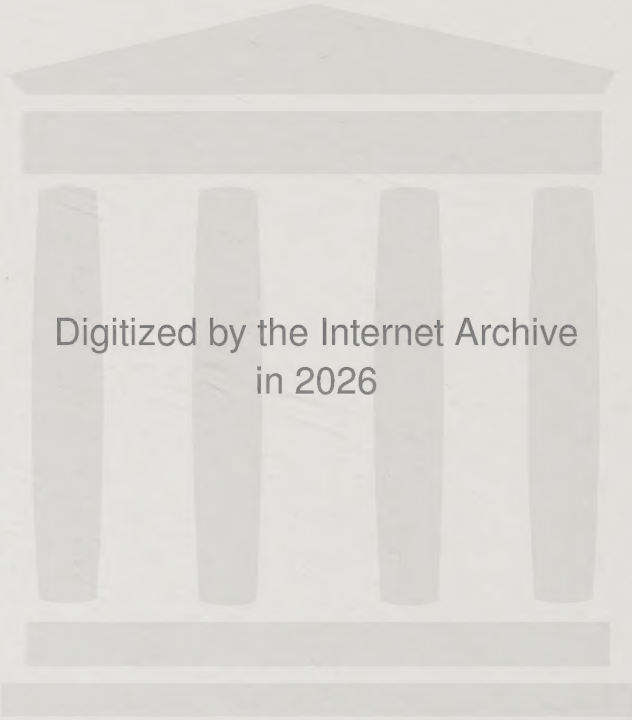


DER EINFLUSS
DES EMOTIONALEN ERREGUNGSPOTENTIALS
VON WÖRTERN AUF DAS VERBALE LERNEN

Elisabeth Schürer-Necker

1981



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548325_0012

DER EINFLUSS DES EMOTIONALEN ERREGUNGSPOTENTIALS
VON WÖRTERN AUF DAS VERBALE LERNEN

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophie
der Ludwig-Maximilians-Universität
zu München

vorgelegt von
Elisabeth Schürer-Necker
aus München

München 1981



Referent: Prof. Dr. Kurt Müller

Korreferent: Prof. Dr. Wolfgang Marx

Tag der letzten mündlichen Prüfung: 23.7.1981

INHALTSVERZEICHNIS

Vorwort	1
---------	---

I THEORETISCHER TEIL

1.	<u>Der Einfluß von Emotionen auf das Gedächtnis</u>	3
1.1.	Überblick über die Literatur	4
1.2.	Annahmen	8
1.3.	Untersuchungen	10
1.3.1.	Fragebogenstudien	11
1.3.2.	Erinnerungsassoziationsversuche	12
1.3.3.	Erinnerungserhebungen	14
1.3.4.	Lernversuche	17
1.4.	Theoretische Positionen	22
1.4.1.	Qualitätsorientierte Positionen	22
1.4.2.	Intensitätsorientierte Positionen	24
2.	<u>Die emotionale Tönung von Wörtern</u>	27
2.1.	Bezeichnungen in der Literatur	30
2.1.1.	Allgemeine Bezeichnungen	30
2.1.2.	Qualitativ orientierte Bezeichnungen	32
2.1.3.	An der Intensität orientierte Bezeichnungen	33
2.2.	Operationalisierungen	34
2.2.1.	Subjektive Einschätzungen (SE)	34
2.2.1.1.	Skalen	37
2.2.1.2.	Reliabilität	42
2.2.1.3.	Probleme der Validität	44
2.2.2.	Physiologische Reaktionen	46
2.2.2.	Exkurs: Die elektrodermale Aktivität	48
2.2.2.1.	Messung der emotionalen Intensität von Wörtern mit Hilfe der EDR	63
2.2.2.2.	Reliabilität	65

2.2.2.3.	Probleme bei der EDR-Messung	67
2.2.2.4.	Validität	69
2.2.3.	Verhaltensmessung	70
2.2.3.1.	Assoziationsverhalten	70
2.2.3.2.	Probleme der Assoziationsparameter	72
2.2.3.3.	Sentogramm	73
2.4.	Das emotionale Erregungspotential	80
3.	<u>Wortcharakteristika und verbales Lernen</u>	82
3.1.	Emotionale Tönung	82
3.1.1.	Qualität der emotionalen Tönung	82
3.1.2.	Intensität der emotionalen Tönung	89
3.1.2.1.	Theorien	89
3.1.2.2.	Untersuchungen	94
3.2.	Bedeutungshaltigkeit (meaningfulness m,m',av,AF)	112
3.2.1.	Operationalisierungen	112
3.2.2.	Einfluß auf das verbale Lernen	116
3.3.	Konkretheit - Abstraktheit (concreteness C, specificity S)	117
3.3.1.	Operationalisierungen	117
3.3.2.	Einfluß auf das verbale Lernen	118
3.4.	Bildhaftigkeit (imagery I, vividness V)	119
3.4.1.	Operationalisierungen	119
3.4.2.	Einfluß auf das verbale Lernen	120
3.5.	Vorkommenshäufigkeit, Vertrautheit (frequency F, familiarity f)	121
3.5.1.	Operationalisierungen	121
3.5.2.	Einfluß auf das verbale Lernen	123
3.6.	Wortart	125
3.7.	Wortlänge	126
3.8.	Interkorrelationen der verschiedenen Wortparameter	127

4.	<u>Einfluß weiterer Faktoren auf das verbale Lernen</u>	134
4.1.	Charakteristika der Wortliste	134
4.1.1.	Beziehungen zwischen den Items	134
4.1.2.	Position	136
4.2.	Situationseinflüsse	137
4.2.1.	Art der Darbietung	137
4.2.2.	Absichtlichkeit vs. Unabsichtlichkeit des Lernens	138
4.2.3.	Länge des Retentionsintervalls	140
4.3.	Erregung und verbales Lernen	141

II EMPIRISCHER TEIL

1.	<u>Fragestellungen und Hypothesen</u>	143
1.1.	Emotionales Erregungspotential	143
1.2.	Emotionales Erregungspotential und Absichtlichkeit des Lernens	145
1.3.	Emotionales Erregungspotential und Zusammensetzung der Lernliste	146
1.4.	Emotionales Erregungspotential und Länge des Retentionsintervalls	146
1.5.	Emotionales Erregungspotential und Geschlecht	148
2.	<u>Operationalisierung der Variablen, Versuchsplan</u>	149
2.1.	Unabhängige Variablen	149
2.2.	Abhängige Variable	151
2.3.	Versuchsplan	151
3.	<u>Versuchsmaterial</u>	152
3.1.	Lernmaterial	152
3.1.1.	Worterhebungen	152
3.1.2.	Lernlisten	156
3.1.2.1.	Willkürliste	156
3.1.2.2.	Clusterliste	157
3.2.	Subjektive Einschätzskalen	161

3.3.	Rekognitionslisten	163
3.4.	Geräte	163
4.	<u>Versuchsdurchführung</u>	164
4.1.	Ort und Zeit	164
4.2.	Versuchsbeschreibung	164
4.3.	Instruktionen	168
4.4.	Stichprobe	169
5.	<u>Auswertung</u>	170
5.1.	Physiologische Daten (PHY)	170
5.2.	Subjektive Einschätzungen (SE)	172
5.3.	Gedächtnisleistung	172
5.3.1.	Free Recall Daten	172
5.3.1.1.	Auswertung nach dem Gruppenurteil	173
5.3.1.2.	Auswertung nach den physiologischen Daten	173
5.3.1.3.	Auswertung nach der subjektiven Einschätzung	174
5.3.2.	Rekognitionsdaten	174
6.	<u>Ergebnisse</u>	175
6.1.	Vorprüfungen	175
6.1.1.	Gesamtgedächtnisleistung	175
6.1.2.	Prüfung des Positionseffekts	178
6.1.3.	Überprüfung des Clusterings	185
6.1.4.	Subjektive Einschätzungen	191
6.2.	Emotionales Erregungspotential und Free Recall	195
6.2.1.	Emotionales Erregungspotential aufgrund des Gruppenurteils	195
6.2.2.	Emotionales Erregungspotential aufgrund der physiologischen Daten	207
6.2.3.	Emotionales Erregungspotential aufgrund der subjektiven Einschätzung	214
6.2.4.	Vergleich der drei Arten des emotionalen Erregungspotentials hinsichtlich der Behaltensleistung	219

6.3.	Emotionales Erregungspotential und Rekognition	224
7.	<u>Zusammenfassung</u>	230
Anhang		232
Literatur		238

Vorwort

In dieser Arbeit soll der Frage nachgegangen werden, welchen Einfluß das emotionale Erregungspotential von Wörtern auf das verbale Lernen hat.

Teil I behandelt die theoretischen Grundlagen, in Teil II soll die von der Autorin durchgeführte Untersuchung dargestellt werden.

Die behandelte Fragestellung ist in dem Bereich angesiedelt, der die Rolle der Emotionen auf die Gedächtnisleistung untersucht. Das 1.Kapitel bringt deshalb einen kurzen Abriß über diesen Forschungsbereich.

Was unter der affektiven Tönung und dem emotionalen Erregungspotential von Wörtern zu verstehen ist, beschreibt das 2.Kapitel.

In Kapitel 3 werden die wichtigsten Untersuchungen und Theorien zum Einfluß der affektiven Tönung von Wörtern auf das verbale Lernen dargestellt und diskutiert. Im selben Abschnitt sollen noch andere Wortcharakteristika erörtert werden, die einen Einfluß auf das verbale Lernen ausüben, da es in der späteren Untersuchung wichtig ist, diese zu kontrollieren.

Weitere Faktoren, die für die Gedächtnisleistung von Bedeutung sind, führt das letzte Kapitel an, soweit sie bei der Untersuchung berücksichtigt werden.

Im experimentellen Teil der Arbeit soll mit Hilfe einer Free Recall Untersuchung sowohl der Einfluß der durch ein Wort ausgelösten physiologischen Aktivierung (gemessen mit der elektrodermalen Reaktion) als auch der Einfluß der beurteilten emotionalen Erregung der Wörter auf das verbale Lernen geprüft werden. Die Untersuchung soll neue Erkenntnisse vermitteln über den differentiellen Einfluß des emotionalen Erregungspotentials von Wörtern bei unterschied-

lichen Lernbedingungen: absichtliches vs. unabsichtliches Lernen, langfristiges vs. kurzfristiges Behalten, Liste mit assoziativ verbundenen Wörtern vs. Liste mit unzusammenhängenden Wörtern. Abschließend werden die physiologischen Reaktionen mit den eingeschätzten Erregungen hinsichtlich der Gedächtnisleistung verglichen.

Mein Dank gilt allen, die mich direkt oder indirekt unterstützt haben, im besonderen Prof. Dr. Kurt Müller und Prof. Dr. Wolfgang Marx.

I THEORETISCHER TEIL

1. Der Einfluß von Emotionen auf das Gedächtnis

Das Hauptaugenmerk gilt in diesem Kapitel, wie auch in der ganzen Arbeit, der Rolle, die Emotionen bei der Einspeicherung von Material in das Gedächtnis spielen. Unberücksichtigt bleibt dagegen der Einfluß von Emotionen zum Zeitpunkt der Reproduktion.

Unsere Alltagserfahrung weist uns immer wieder darauf hin, daß es eine Beziehung zwischen Emotionen und Gedächtnis geben muß. Es hat den Anschein, daß man sich hauptsächlich an Ereignisse erinnert, die mit stärkeren Emotionen verbunden waren. Dinge, die einen nicht berühren, werden schnell vergessen. Manchmal hat man auch den Eindruck, daß von Erlebnissen die unangenehmen Seiten immer mehr verblasen, und man sich vor allem an die positiven Aspekte erinnert. Dieser Umstand kann eine Verklärung der Vergangenheit bewirken, was sich in der häufig zu hörenden Redensart 'die gute alte Zeit' äußert. Man erlebt aber auch das Gegenteil: manche Menschen erinnern sich fast nur an negative Ereignisse.

Ein anderer Hinweis darauf, daß Emotionen eine wichtige Rolle spielen, ist die Tatsache, daß schockierende Erlebnisse zu einer 'Erinnerungslöschung' führen können. Als Beispiel soll ein Artikel aus der Süddeutschen Zeitung vom 30.5.1980 (S. 43) dienen, der die Überschrift trägt: "Anblick der Pistole löscht die Erinnerung". Ein Phantomzeichner der Kriminalpolizei teilt seine Erfahrung aus 35 Dienstjahren mit: Bei Raub, Sexualdelikten oder Banküberfällen ist es meist der Anblick der vorgehaltenen Tatwaffe, der Opfer und Zeugen so schockiert, daß sie keine detaillierten Aussagen über den Täter machen können.*

* Hier bleibt unklar, ob das Opfer vor Schreck nur auf die Tatwaffe starrt und deshalb den Täter nicht beachtet. Diese Einschränkung des Blickfeldes würde dann eine reduzierte Wahrnehmung nach sich ziehen, der Bedrohte nimmt seine Umwelt nicht im gewohnten Umfang wahr und kann deshalb später

1.1. Überblick über die Literatur

Wenden wir uns nun dem Bereich der Psychologie zu. Was wird hier zu diesem Thema ausgesagt? Es gibt, vor allem in neuerer Zeit, nur vereinzelt Bücher, die einen Überblick über das zu behandelnde Thema vermitteln. Hier ist an erster Stelle RAPAPORT mit seinem Buch "Gefühl und Erinnerung" (1950/1977) zu nennen, als weitere Arbeit das Buch von DUTTA und KANUNGO "Affect and Memory: A Reformulation" (1975), die sich aber teilweise stark an RAPAPORT orientieren. Öfters finden sich in Zeitschriftenartikeln Abrisse zu diesem Gebiet. Besonders in der ersten Hälfte unseres Jahrhunderts, in der viele Untersuchungen zum Einfluß der Gefühle auf das Gedächtnis durchgeführt wurden, entstanden einige Überblicksartikel und Bücher. Diese werden nicht gesondert hier aufgeführt. Ausführliche Literaturangaben können bei RAPAPORT, DUTTA & KANUNGO oder auch McGEOCH & IRION (1952) nachgelesen werden. Der noch unveröffentlichte Überblick von BOCK (1978 a) in seiner Antrittsvorlesung mit dem Thema "Angenehme und unangenehme Erfahrungen aus gedächtnispsychologischer Sicht" ist eine der jüngsten Darstellungen. Er bezieht auch die Untersuchungen zum Einfluß des Arousal* auf das Gedächtnis aus den Sechziger und Siebziger Jahren mit ein, die bei DUTTA & KANUNGO (1975) nicht erwähnt werden.

keine genauen Auskünfte geben. Andererseits wäre es auch möglich, daß das Opfer zwar alle wichtigen Faktoren registriert, sie aber bei einer späteren Befragung nicht wiedergeben kann. Zu klären bliebe hier, ob starke Emotionen schon die Wahrnehmung beeinträchtigen oder erst die Speicherung oder noch später die Reproduktion. Mit dem Blickregistrierungsgerät ließe sich untersuchen, ob starke Emotionen die Wahrnehmung stören. Träfe dies zu, so müßte das Problem wohl eher im Rahmen der Wahrnehmungspsychologie behandelt werden.

* Arousal = Erregung, Aktivierung

Auf dieses Teilgebiet (Arousal und Gedächtnis) wird unter Punkt 3.1.2. der Arbeit näher eingegangen.

Eine Übersicht über Untersuchungen zur Verdrängungstheorie von FREUD stammt von ZELLER (1950 a) und eine neuere von HOLMES (1974).

Kurz gestreift wird das Thema in mehreren Büchern zur Allgemeinen Psychologie, vor allem der Lern- und Gedächtnispsychologie, und zur Experimentellen Psychologie, z.B. von McGEEOCH & IRION (1952), OSGOOD (1964, Kap. 'Motivation and Memory'), TRAXEL (1963, in MEILI & ROHRACHER) usw.

Es existiert aber auch eine Reihe von Abhandlungen über das Gedächtnis und Gedächtnistheorien, die den Einfluß von Emotionen außer acht lassen. BOCK (1978 a) bemerkt, daß in den neueren zusammenfassenden Darstellungen der gegenwärtigen Gedächtnispsychologie (TULVIN & MADIGAN 1970, POSTMAN 1975 oder PETERSON 1977) nicht zu diesem Problem Stellung genommen wird. Es wäre notwendig zu prüfen, warum Emotionen bei den Untersuchungen und Theorien mancher Gedächtnisforscher keine Rolle spielen. Liegt es daran, daß sie Gefühle für unwichtig erachten? Halten sie das Vorhandensein von Emotionen für eine Selbstverständlichkeit, die nicht extra untersucht zu werden braucht, da ihre Funktion klar ist? Oder wird die Erforschung der Gefühle als zu schwierig angesehen, als daß sie realisiert werden könnte.

Die Emotionen sind aber keineswegs ganz v e r g e s s e n worden. Es gibt auch in neuerer Zeit Wissenschaftler, die die Rolle von Gefühlen bei der Entwicklung ihrer Gedächtnismodelle berücksichtigen, z.B. OTT & MATTHIES (1978).

Bei der Durchsicht der Literatur stellt man fest, daß ausführliche Behandlungen des Themas äußerst selten sind. Aus den kurzen Ausführungen zu diesem Bereich läßt sich herauslesen, daß Emotionen als wichtiger Faktor für das Gedächtnis angesehen werden. Schon BACON (1870) bemerkte (laut RAPAPORT 1977, S. 24), daß Tatsachen, die mit starken Gefühlen (wie Beschämung, Furcht, Bewunderung) verbunden sind, leichter behalten werden als gleichgültige Tatsachen. WECHSLER schreibt (1925, S.132):

One of the most important psychological aspects of

emotion is its effect upon the mnesic processes. It is a matter of common experience that an event which aroused us emotionally at the time of its occurrence is more likely to be remembered than one to which we were entirely indifferent. On the other hand, acts during outbursts of passion and in great excitement are likely either to entirely escape our notice or to be quickly forgotten. In general, there is a tendency to remember the pleasant and to forget quickly disagreeable experiences.

FRÜBES (1929, S.287) zitiert KORTSEN "Erlebnisse werden behalten, wenn starke Reize dabei eine Rolle spielten, oder wenn sie bedeutungsvoll, von lebhaften Gemütsbewegungen begleitet waren" und PETERS und PICK, die beide die Tendenz der Unterdrückung unangenehmer Erlebnisse erwähnen.

Im "Wörterbuch der Psychologie" (CLAUSS, Hrsg., 1976, S.191) liest man: "Die Speicherung des G.materials hängt von dessen biologischer Bedeutung bzw. der affektiven Wertigkeit .. ab: Bedeutsames wird rascher und über längere Zeit gespeichert als Bedeutungsloses."

Die angeführten Zitate lassen keinen Zweifel daran, daß alle Autoren, die sich zu dem Thema äußerten, übereinstimmend der Ansicht sind, daß Emotionen eine wichtige Rolle bzgl. des Gedächtnisses spielen. Die meisten Äußerungen bewegen sich auf einem sehr allgemeinen und oberflächlichen Niveau, sie geben vor allem Alltagserfahrungen wieder. Neben der Betonung, daß Emotionen eine Rolle spielen, ist nur sehr allgemein von fördernden und von hemmenden Einflüssen die Rede. Stark emotional gefärbte Ereignisse werden besser behalten, zu starke Emotionen können aber auch hinderlich sein. Neben diesen nur auf das Vorhandensein und die Stärke des Gefühls gerichteten Aussagen finden sich auch solche, bei denen die positive oder negative Färbung von Emotionen als Grund für das Behalten angenehmer und das Vergessen unangenehmer Dinge angesehen wird.

Bei den ausführlicheren Behandlungen offenbaren sich die Schwierigkeiten dieses Forschungsbereichs. Abgesehen von dem Resümee, daß Emotionen eine Rolle spielen, lassen sich meistens keine dezidierten Aussagen feststellen. Die referierten Untersuchungen zeigen oft gegensätzliche Ergebnisse und leiden an methodischen Unzulänglichkeiten.

Warum das Verhältnis zwischen Emotionen und Gedächtnis bis heute noch nicht genau geklärt ist, hat m. E. verschiedene Ursachen. Es existiert noch keine allgemeingültige Definition* von 'Emotion' und von 'Gedächtnis'. Beide Bereiche sind sehr umfangreich, diffus und komplex. Weiter wird man mit Operationalisierungsschwierigkeiten vor allem auf Seiten der Emotionen konfrontiert und schließlich fehlt eine einheitliche, explizite Theorie.

Betrachtet man die Untersuchungen, so zeigt sich, daß unter 'Gedächtnis' vor allem das bewußte Sich-Erinnern verstanden wird. Außer acht bleiben eventuell unbewußt wirkende, dem Individuum nie bewußt verfügbare Gedächtnisinhalte. Von meiner Seite erfolgt in dieser Arbeit die gleiche Einschränkung des 'Gedächtnis' auf die Reproduktion und Rekognition, da sie methodisch leichter zu erfassen sind, und es nötig ist, die Frage bzgl. des Einflusses der Emotionen auf das Gedächtnis zu konkretisieren, um sie empirisch angehen zu können.

Bzgl. der Emotionen wurden hauptsächlich zwei Aspekte bearbeitet: 1) die Qualität, wobei meist nur die grobe Klassifizierung 'positiv' ('angenehm', 'lustbetont') vs. 'negativ' ('unangenehm', 'unlustbetont') und vereinzelt noch 'neutral' (oder 'indifferent') herangezogen wurde, und 2) die Intensität.

Bei den untersuchten 'Emotionen' handelte es sich in vielen Fällen nur um beurteilte Gefühle. Die Einschätzung

* Definitionsversuche können z.B. bei RAPAPORT (1977) nachgelesen werden.

wurde in einigen Untersuchungen vom Forscher vorgenommen, meistens beurteilte aber die Vp selber die Gefühle.

Aber auch mit Hilfe physiologischer Messungen wurde der Versuch unternommen, Gefühle zu erfassen. Schon früh hatte man erkannt, daß Gefühle von physiologischen Veränderungen begleitet oder, je nach theoretischer Auffassung, ausgelöst werden (JAMES-LANGE Theorie: "wir sind traurig, weil wir weinen"). Gemessen werden vor allem Hautwiderstand, Pulsfrequenz, Atmung, Blutdruck, hirnelektrische Aktivität.

Bis heute ist es noch nicht gelungen, die physiologischen Korrelate der Gefühle in den Griff zu bekommen. Man hat dabei mit mehreren Schwierigkeiten zu kämpfen: die oft geringen Interkorrelationen der verschiedenen physiologischen Variablen (BOTTENBERG, 1972), die großen interindividuellen Unterschiede in den physiologischen Reaktionen (unterschiedliche Aktivierung der einzelnen Systeme bei verschiedenen Vpn), die ähnlichen physiologischen Erscheinungsformen unterschiedlicher Emotionen. Aus den physiologischen Erscheinungsmustern lassen sich keine genau definierten Emotionen ablesen. Feststellen läßt sich aber die Stärke der Aktivierung. Die Bestimmung der Qualität (positiv oder negativ) ist bei den meisten Indikatoren nicht möglich.

1.2. Annahmen

Die Überlegungen zum Einfluß von Emotionen auf das Gedächtnis erbrachten zwei Hauptannahmen:

- 1) Die Qualität der Emotionen bestimmt, was im Gedächtnis haften bleibt.
- 2) Nur die Intensität, unabhängig von der Art des Gefühls, ist entscheidend.

Folgt man den Vertretern der ersten Richtung, so hieße das, daß angenehme Inhalte gut im Gedächtnis behalten werden, während die Tendenz zum Vergessen bei unangenehmen Inhalten sehr viel größer ist.

Die 2. Annahme wurde meist so formuliert: Intensivere (gefühlsstarke) Erlebnisse werden besser erinnert als schwache. Wie dieses Verhältnis genau aussieht, wurde nicht spezifiziert. Man könnte annehmen, daß hier eine positiv lineare Beziehung vorliegt: je intensiver die ausgelösten Gefühle, desto besser das Behalten. Eine andere mögliche Beziehung könnte eine kurvilineare sein, wie z.B. die umgekehrte U-Funktion zwischen Erregung und Leistung oder Motivation und Leistung. Gefühle mittlerer Intensität führen zur besten Behaltensleistung, schwache und starke Gefühle sind mit geringerer Behaltensleistung verbunden.

Neben diesen beiden Positionen, die jeweils nur einen Aspekt berücksichtigen, kann man sich aber auch noch eine 3. Sichtweise vorstellen, die beide Aspekte miteinbezieht, daß sowohl die Qualität als auch die Intensität eine Rolle spielen. Das Zusammenwirken beider Faktoren bestimmt die Behaltensleistung. Beispielsweise wären folgende Annahmen möglich: Bei unangenehmen Ereignissen liegt eine umgekehrte U-Funktion zwischen Intensität und Behalten vor, bei angenehmen Ereignissen eine positiv lineare Beziehung. Oder: unangenehme Erlebnisse werden nur behalten, wenn sie sehr intensiv sind, angenehme Erlebnisse dagegen bleiben immer haften, auch wenn sie äußerst schwach sind. Diese 3. Sichtweise könnte man der Äußerung von PETERS (1911) entnehmen, der vermutet, daß die erinnerten Unlusterlebnisse eine stärkere Bedeutung gehabt hätten als die erinnerten lustbetonten Erlebnisse und "... daß die unlustbetonten Erlebnisse von geringerer Bedeutung weniger Chancen haben, erinnert zu werden, als die lustbetonten Erlebnisse von gleicher Bedeutung" (aus RAPAPORT, 1977, S.77), wenn man 'Bedeutung' mit 'Stärke' des Gefühls gleichsetzt. In der Literatur finden sich zur 3. Sichtweise keine expliziten Aussagen über mögliche Interaktionen und Wechselwirkungen der beiden Faktoren Qualität und Intensität.

1.3. Untersuchungen

Die ersten Untersuchungen, die sich mit dem Thema befaßten, vor allem zu Beginn unseres Jahrhunderts, beschäftigten sich fast ausschließlich mit der 1. Annahme. Man wollte feststellen, ob vor allem angenehme Erfahrungen im Gedächtnis haften bleiben und ob unangenehme Erlebnisse vergessen werden. Mit verursacht waren viele dieser Untersuchungen von der Verdrängungstheorie von FREUD. Sie behauptet, daß konfliktträchtige, angsterzeugende und peinliche Erfahrungen oder Ereignisse verdrängt werden. Laut RAPAPORT wurde diese Theorie von vielen Forschern falsch aufgefaßt. Die Verdrängungstheorie besage, daß Vergessen die "Folge der Tendenz, 'die Erweckung von Unlust durch Erinnern zu vermeiden'" ist, und nicht, daß allgemein die Neigung besteht, unangenehme Dinge zu vergessen (1977, S.180). Letzteres wurde zwar auch von FREUD geäußert, RAPAPORT kritisiert dies aber als irreführende Formulierung.

Neben der Verdrängungstheorie von FREUD sieht RAPAPORT noch zwei weitere Faktoren, die diese Untersuchungen beeinflussen haben. Zum einen erwähnt er die Fin-de siècle Stimmung und die pessimistische Philosophie SCHOPENHAUERS, wodurch die Frage aufgeworfen wurde, ob der Mensch im wesentlichen optimistisch oder pessimistisch sei. Das Problem, ob angenehme oder unangenehme Erfahrungen besser behalten werden, wurde dadurch für die Gedächtnispsychologie relevant.

Als dritten Einfluß nennt er die 'Lust-Unlust-Theorie' des Lernens, die von SPENCER (1873) und BAIN (1868) aufgestellt und ausgearbeitet wurde. "Spencer definierte Lust als 'ein Gefühl, das wir ins Bewußtsein bringen möchten' und Unlust als 'ein Gefühl, das wir aus dem Bewußtsein drängen möchten' und lehrte infolgedessen, Lusterlebnisse würden häufig erinnert und Unlusterlebnisse würden vergessen." (1977, S.65/66)

Die ausschließliche Konzentration auf die Qualität der Gefühle nahm mit der Zeit immer mehr ab. Das Interesse verlagerte sich zunehmend auf die Intensität der Gefühle, da die

Untersuchungen zur Qualität keine schlüssigen Ergebnisse erbrachten. In jüngster Zeit gibt es kaum noch Untersuchungen zum Einfluß der Gefühlsqualität, es werden vor allem Untersuchungen zum Intensitätsaspekt durchgeführt, die in den letzten Jahren unter der Bezeichnung 'Arousal' laufen.

Man kann die verschiedenen Untersuchungen in vier Gruppen einteilen:

- 1) Fragebogenstudien (Hier wird nur die Meinung über das Gedächtnis erfragt.)
- 2) Erinnerungsassoziationsversuche (Auf Reizwörter soll mit Erlebnissen reagiert werden. Die affektive Tönung der Erlebnisse wird anschließend analysiert.)
- 3) Erheben von Erinnerungen an Erlebnisse
- 4) Lernversuche.

Im folgenden Abschnitt soll ein kurzer Blick auf eine Reihe bisher durchgeführter Untersuchungen geworfen werden.*

1.3.1. Fragebogenstudien

Mit Hilfe von Fragebogenstudien wollte man erfahren, ob die Menschen sich besser an angenehme oder an unangenehme Erlebnisse in ihrem Leben erinnern. Die erste Studie stammte von COLGROVE (1899, 1900). Er befragte Gruppen unterschiedlichen Alters und verschiedener Rassenzugehörigkeit, erhielt aber keine eindeutigen Ergebnisse, da die Antworten stark variierten.

KOWALEWSKI (1904) befragte Kinder und Jugendliche im Alter von 10 bis 13 Jahren und stellte zwei Typen fest, die er als 'Erinnerungsoptimisten' und 'Erinnerungspessimisten' bezeichnete. Die einen glaubten, daß sie sich besser an die ange-

* Von den Untersuchungen in der ersten Hälfte des Jahrhunderts konnte nur ein Teil von der Autorin im Original gelesen werden. Die Literaturangaben der Untersuchungen, die nur in der Sekundärliteratur gelesen werden konnten, finden sich bei RAPAPORT (1977), DUTTA & KANUNGO (1975), McGEOCH & IRION (1952) und OSGOOD (1964).

nehmen Dinge erinnern, sie machten die Mehrzahl seiner Vpn aus. Die anderen vermuteten, daß sie sich besser die unangenehmen Dinge merken.

SUSUKITA (1934) fand in seiner Studie, daß der Gedächtnisoptimismus unter schlechten soziokulturellen Bedingungen vorherrschte, der Gedächtnispessimismus unter besseren soziokulturellen Bedingungen. Ebenso stellte er fest, daß mit dem Alter der Gedächtnispessimismus zunahm. Da er nur Kinder untersuchte, läßt sich sein Ergebnis nicht ohne weiteres auf Erwachsene übertragen. Andere Autoren, z.B. W.PETERS (1911) vermuten eine mit dem Alter zunehmende Tendenz, sich vor allem an die angenehmen Dinge zu erinnern.

In den Fragebogenstudien wurde nicht tatsächlich das Gedächtnis geprüft, sondern nur die Ansicht der Vpn über ihr Gedächtnis erfragt. Diese Methode ist nicht geeignet, Aufschluß über den Einfluß von Emotionen auf das Gedächtnis zu erhalten, da hier nicht das tatsächliche Erinnerungsverhalten erfaßt wird.

1.3.2. Erinnerungsassoziationsversuche

Erinnerungsassoziationsversuche wurden ebenfalls herangezogen, um das Verhältnis 'Emotion - Gedächtnis' zu klären. Der Vp werden Reizwörter vorgegeben, auf die sie mit Erlebnissen reagieren soll. Sowie ihr ein Erlebnis einfällt, antwortet sie mit 'ja'. Dann berichtet sie das Erlebnis und beantwortet einige Fragen über die Gefühlsbetonung.

Zu den klassischen Arbeiten auf diesem Gebiet zählt die Untersuchung von W.PETERS (1914). Er erhob bei 8 Vpn ca. 900 Erinnerungserlebnisse. Die Vpn sollten bei jedem Erlebnis u.a. angeben, ob es gefühlsbetont war und wenn ja, welchen Gefühlston es hatte, wie lange das Erlebnis zurücklag und wie oft das Erlebnis erinnert wurde. Es zeigte sich, daß 80% der assoziierten Erlebnisse gefühlsbetont waren und von diesen 65% lustbetont (A)*, 30% unlustbetont (U)* und 5%

* Im folgenden Text bedeutet A 'angenehm', 'positiv', 'lust-

gemischt (G). Die länger zurückliegenden Erlebnisse waren häufiger lustbetont. Unter den häufig erinnerten Erlebnissen war die Mehrzahl von Unlustgefühlen geprägt. Um dieses Ergebnis zu erklären, verwies PETERS auf die Aussagen der Vpn, die ihre negativ getönten Erlebnisse sehr oft als für sie wichtig ansahen, ihre A-Erlebnisse dagegen häufiger den unwichtigen zuordneten. PETERS stellte weiter fest, daß bei U-Erlebnissen die Reaktionszeit länger war als bei A- oder N-Erlebnissen.*

Der problematische Punkt bei dieser Methode ist die Auswahl der Reizwörter. Bietet man der Vp hauptsächlich negativ getönte Reizwörter an, wie z.B. 'Tadel'. 'Schmerz', 'Trauer' usw., so wird man vermutlich eine andere Verteilung mit Überwiegen der U-Erlebnisse erhalten. In Assoziationsversuchen (POLLIO 1963 a und GEER & MOLLENAUER 1964; zit. nach CRAMER 1968) wurde gezeigt, daß die Reaktionswörter meistens die gleiche emotionale Tönung aufweisen wie das Reizwort. Wahrscheinlich läßt sich dieser Befund auch auf assoziierte Erlebnisse übertragen.

Dadurch, daß nur das erste Erlebnis erfaßt wurde, lassen sich keine Aussagen über das Gedächtnis an sich machen. Es könnte sein, daß die später einfallenden Erlebnisse einen eher negativen Charakter aufweisen, da A-Erlebnisse in der Untersuchung von PETERS kürzere Reaktionszeiten hatten.

betont' (englisch meist 'pleasant' oder 'good'); U 'unangenehm', 'negativ', 'unlustbetont' ('unpleasant' oder 'bad'). Für 'neutral' oder 'indifferent' wird N verwendet.

* Auch die momentane Stimmung der Vp kann eine Rolle spielen. TEASDALE & FOGARTY (1979) vermittelten vor Beginn des Versuchs einem Teil ihrer Vpn eine positive Stimmung, den Rest brachten sie in depressive Verfassung. Die deprimierten Vpn brauchten im Vergleich zu den positiv gestimmten Vpn länger, bis ihnen A-Erlebnisse einfielen (relativ zu den U-Erlebnissen).

1.3.3. Erinnerungserhebungen

Bei diesen Studien versuchte man, die Erinnerung an Erlebnisse des realen Lebens direkt zu prüfen. Die Vpn sollten entweder eine bestimmte Anzahl von Erlebnissen berichten oder Erlebnisse, die einen bestimmten Abschnitt ihres Lebens betrafen. Die einfachsten Studien erhoben nur Erinnerungen an Erlebnisse, ohne zu kontrollieren, welche Erlebnisse die Vp tatsächlich hatte, und stellten den Anteil von angenehmen und von unangenehmen Erlebnissen fest. KOWALEWSKI (1908) befragte Schulkinder über ihre Erlebnisse an einem Feiertag einen Tag danach und 10 Tage später. Beim ersten Mal gaben 62% der Kinder, bei der zweiten Prüfung 61,4% der Kinder mehr angenehme als unangenehme Erlebnisse wieder. KOWALEWSKI folgerte daraus, daß in der Erinnerung der meisten Menschen die angenehmen Erlebnisse besser haften bleiben. Daß in den Berichten der Vpn die angenehmen Ereignisse überwogen, kann auch daran gelegen haben, daß sie mehr angenehme Dinge erlebten. Hier existiert implizit die Annahme eines Verhältnis von 50% A- zu 50% U-Erfahrungen im Leben.

Ein gegenteiliges Ergebnis erhielt GORDON (1928), die ihre Vpn über deren früheste Kindheitserlebnisse befragte und dabei feststellte, daß die meisten unangenehmer Art waren.

HENDERSON (1911) ließ sich 100 Erlebnisse berichten. Neben der qualitativen Beurteilung verlangte er zusätzlich eine Einschätzung der Intensität. Laut RAPAPORT (1977, S.100) ist HENDERSON der erste V1, der auch die Stärke der Gefühle miteinbezog. Es fanden sich mehr A- als U-Erlebnisse und mehr U- als N-Erlebnisse; ebenso gab es mehr starke als oberflächliche Erlebnisse. HENDERSON interpretierte seine Ergebnisse anders als KOWALEWSKI, da nach seiner Meinung nicht klar ist, wie das Verhältnis von A- und U-Erlebnissen im Alltagsleben aussieht. Hatte die Vp gleich viele A- wie U-Erlebnisse oder hatte sie mehr A-Erlebnisse?

Da es bei diesen Untersuchungen an einer Kontrolle fehlte, welcher Art die gesamten Erlebnisse waren, die der Befragte hatte, konnten keine Aussagen getroffen werden über eine angebliche Bevorzugung angenehmer Erlebnisse in der Erinnerung. Die erinnerten Erlebnisse könnten die gleichen Verhältnisse widerspiegeln wie die erfahrenen Erlebnisse. Weil diese Daten nicht vorhanden waren, konnte man wenig mit den Ergebnissen anfangen.

Fragen muß man sich auch, inwieweit die Vpn bereit sind, alle Erinnerungen tatsächlich mitzuteilen und inwieweit sie unangenehme Erlebnisse bewußt verschweigen. Weiter kann die Artikulationsfähigkeit eine Rolle spielen; denkbar wäre, daß die Vp gar nicht in der Lage ist, bestimmte Erlebnisse zu verbalisieren.

In den methodisch besseren Studien bemühte man sich, zusätzlich die Erlebnisse zu registrieren. FLUEGEL (1925) ließ 9 Vpn einen Monat lang jede Stunde ihre Erlebnisse aufschreiben und dabei die Stärke, Dauer und Qualität der Gefühle festhalten. Es wurden mehr A-Erlebnisse notiert. FLUEGEL zog daraus den Schluß, "daß 'Lust einen erheblich größeren Teil des menschlichen Lebens ausfüllt als Unlust' (100, S.328 f)." (aus RAPAPORT 1977, S.101). Bei diesem Ergebnis kann man jedoch Zweifel anmelden: Inwieweit wurde durch die bewußte Beobachtung die Wahrnehmung der Erlebnisse verfälscht? Weiter ist die Auswahl der Vpn kritisierbar. Da es sich nur um 9 Vpn handelte, die auch noch alle die Zeit hatten, stündlich einen Bericht zu schreiben, scheint es sich um eine sehr selektive Stichprobe zu handeln, und es ist deshalb vor einer Generalisierung auf andere Vpn zu warnen.

Interessant ist der Versuch von KOCH (1930), die 76 Studenten im Laufe von 10 Benotungen mündlicher Leistungen im Anschluß an die Benotung um die Einschätzung ihrer Gefühlsreaktionen bat und 5 Wochen später ihre Erinnerung an die 10 Noten prüfte. Die Vpn erinnerten sich vor allem an die guten

Noten, aber auch an die schlechten Noten, wenn diese stark gefühlbetont beurteilt worden waren.

MELTZER (1930) ließ 132 Studenten ihre Erlebnisse in den Weihnachtsferien bzgl. der Qualität (A - U - N) und der Intensität (auf einer 5-Punkte Skala) einschätzen und prüfte 6 Wochen später die Erinnerung daran. Von den angenehmen Erlebnissen wurden 42,6% vergessen, von den unangenehmen 59,73%. Viele Vpn merkten sich vor allem die angenehmen Ereignisse, es gab aber auch Vpn, die mehr A- als U-Erlebnisse vergaßen.

Einen Vergleich von Tagebuchberichten oder Erlebnisaufzeichnungen mit den ein paar Tage bis Wochen später stattfindenden Gedächtnisprüfungen (Erinnern der registrierten Erlebnisse) führten auch JERSILD (1931), CASON (1932), THOMSON (1930), MENZIES (1935), SUSUKITA (1935), WATERS & LEEPER (1936), O'KELLEY & STECKLE (1940) durch. Meistens wurden mehr A-Erlebnisse berichtet. Bzgl. der späteren Erinnerungsleistung kam es zu sich widersprechenden Ergebnissen. Manche Forscher stellten ein stärkeres Überwiegen von A-Erinnerungen fest, manche fanden die gleichen Verhältnisse wie beim Bericht vor und manche konstatierten ein rascheres Vergessen der angenehmen Erlebnisse. MENZIES (1935) und CASON (1932) zogen aus ihren Daten den Schluß, daß nicht die Qualität, sondern die Intensität der Gefühlsbetonung mit dem Gedächtnis zusammenhängt. Siehe z.B. MENZIES (1935, S.278)

Thus, frequency and extent of recall for ordinary life experiences appear to be correlated with INTENSITY of feeling-tone rather than QUALITY of feeling-tone.

MENZIES, der mehrmals in Abständen die Erinnerung prüfte und dabei jedesmal die Erlebnisse neu beurteilen ließ, bemerkte außerdem, daß die von der Vp beurteilte Stärke der Affektbetonung mit der Zeit abnahm. Zu fragen wäre, ob nicht ein direkter Zusammenhang zwischen dem Nachlassen der Gefühlsintensität und dem Vergessen der Erlebnisse im Lauf der Zeit besteht.

Da die Versuche bzgl. der Qualität der Gefühle keine konsistenten Ergebnisse erbrachten, wandte man sich immer mehr der Betrachtung des Intensitätsaspekts zu. Es zeigte sich meist, daß Erlebnisse, die mit starken Gefühlen verbunden waren, besser behalten wurden als schwache und indifferente Erlebnisse.

Eine Untersuchung neueren Datums ist die von TURNER & BARLOW (1951). Die Vpn sollten ihre A- und U-Erlebnisse nach der Stärke des Gefühls in eine Rangreihe bringen. 14 Tage später fand die Behaltensprüfung an die Erlebnisse statt. Ihre Ergebnisse machen deutlich, daß die Behaltensleistung allein von der Emotionsstärke abhängt, unabhängig von der Qualität. Je intensiver die Erlebnisse eingestuft wurden, desto besser wurden sie behalten.

1.3.4. Lernversuche

Dabei handelt es sich um eine ganz andere Art von Versuchen, bei denen das zu behaltende Material den Vpn erst dargeboten und anschließend die Erinnerung daran geprüft wird. Die wesentlichen Unterschiede zu den zuvor besprochenen Versuchen bestehen darin, daß diese Lern- und Gedächtnisversuche im Labor stattfinden und daß das von den Vpn zu erinnernde Material vom V1 eingeführt und manipuliert wird. Diese Art von Versuchen, die keine Erlebnisse, sondern isolierte Stimuli wie Farben, Gerüche, Töne, Bilder, Wörter usw. verwenden, wurde von Forschern, die mit Erlebnissen aus dem Alltagsleben arbeiteten als zu künstlich und lebensfremd kritisiert. Im Labor könne man keine Emotionen wie im wirklichen Leben hervorrufen. Andererseits gestatten aber erst diese Lernversuche eine Manipulation und Kontrolle des zu behaltenden Materials, was ja der große Nachteil bei den Erlebnisberichten war.

Die Reize werden dabei entweder vom V1 hinsichtlich ihrer emotionalen Tönung beurteilt oder von einer Gruppe von Personen, die nicht an den Gedächtnisversuchen teilnehmen oder von der Vp selbst, mit der der Gedächtnisversuch durchge-

führt wird. Bei einigen Experimenten werden physiologische Messungen mit herangezogen, um anhand der gewonnenen Werte die Gefühlsintensität zu bestimmen.

Die Lernversuche umfassen sowohl Versuche, die mit unabsichtlichem Lernen arbeiten, als auch solche, bei denen absichtliches Lernen verwendet wird. Die Untersuchungen zum inzidentellen Lernen machen die Mehrzahl der Fälle aus.

Die durchgeführten Versuche lassen sich in drei Gruppen unterteilen:

- a) Das Behalten emotional getönter Reize wird mit der Methode des Free Recall oder der Recognition untersucht.
- b) Die Erinnerung an Material wird geprüft, das mit emotionalen Stimuli gekoppelt wurde.
- c) Der Einfluß emotionaler Zusatzreizung während oder nach der Darbietung des Behaltensmaterials wird untersucht.

zu a) Als Lernmaterial dienten entweder Farben, Bilder oder Wörter. GORDON (1905) zeigte ihren Vpn farbige Muster verbunden mit der Anweisung, diese hinsichtlich A, U oder N zu beurteilen. Unmittelbar nach Versuchsende und drei Wochen später wurden die Vpn gebeten, die farbigen Muster anzugeben, an die sie sich erinnern konnten. Dabei stellte GORDON keine Bevorzugung von A, U oder N fest.

TAITs Untersuchung (1913) über das sofortige Wiedererkennen von A-, U- und N-Farben erbrachte eine bessere Leistung für A-Farben. Am schlechtesten wurden die N-Farben wiedererkannt. (A: 63,4%; U: 47,2%; N: 27,3%)

Wiedererkennungsprüfungen wurden auch mit Bildern (Portraits, Kunstpostkarten) durchgeführt. A.PETERS (1917, zit. nach FRÖBES 1929) stellte fest, daß lustbetonte Bilder besser wiedererkannt werden als unlustbetonte und diese besser als indifferente. Ebenso werden stärker gefühlsbetonte besser wiedererkannt als die weniger gefühlsbetonten.

Mit Wörtern als Stimuli fand eine große Anzahl von Untersuchungen statt, die ausführlicher unter Punkt 3.1. der

Arbeit besprochen werden. Nur soviel sei hier schon gesagt, daß die Untersuchungen zur Qualität inkonsistente Ergebnisse erbrachten*, die Untersuchungen zur Intensität dagegen vielversprechender ausfielen. Die Daten weisen in die Richtung, daß die Gefühlsintensität die wichtigere Variable ist und daß die Qualität vermutlich keine oder eine sehr untergeordnete Rolle spielt.

zu b) Bei den Paarbildungsversuchen wurden meistens Gerüche als emotionale Stimuli verwendet, teilweise aber auch Farben und Tonhöhen. Sie wurden gekoppelt mit Bildern, sinnlosen Silben, Namen oder Zahlen.

RATLIFF (1938) stellte bei der Gedächtnisleistung eine Überlegenheit der mit U-Gerüchen verbundenen Zahlen gegenüber den mit A-Gerüchen verknüpften fest. Farben und Tonhöhen als Stimuli bewirkten dagegen ein besseres Behalten der Zahlen, die mit A-Farben oder A-Tonhöhen gekoppelt waren. RATLIFF nahm daher an, daß der Verdrängungsmechanismus aufgrund seiner vermuteten Lokalisation im Thalamus nur beim Hören und Sehen, nicht aber beim Riechen wirksam würde.

Andere Studien stellten keinen Unterschied im Behalten von Material, das mit A-Stimuli und Material, das mit U-Stimuli verbunden war, fest (Z.B. GORDON 1928).

WÜRDEMAN (1926) untersuchte nicht den Einfluß der Qualität von A und U, sondern den Einfluß der Stärke der Gefühle, da er bezweifelte, daß jedes Gefühl entweder nur eine angenehme oder nur eine unangenehme Qualität aufweist. Bei seiner Untersuchung kombinierte er Figuren mit Gerüchen und kam zu dem Ergebnis, daß Sinneseindrücke, bei denen Tiefe oder Intensität des Gefühls groß war, länger in der Erinnerung haften bleiben als solche mit

* In den Versuchen, die nur die Qualität prüften, wurde die Intensität nicht kontrolliert. Es könnte leicht sein, daß in diesen Versuchen A und U unterschiedlich stark ausgeprägt waren und dadurch die unterschiedlichen Ergebnisse verursacht haben.

schwachen Gefühlen.

zu c) Hier dienten als störende Einschübe Gerüche, laute Geräusche (z.B. Pistolenschüsse), Lichtblitze, emotionsauslösende Bilder, experimentell erzeugte Schwindelgefühle, unangenehme Vorstellungen, elektrische Schläge und induzierte Mißerfolgs- oder Erfolgserlebnisse. Bei dem zu behaltenden Material handelte es sich meist um neutrale Wörter oder sinnlose Silben.

TAIT (1913) erzielte mit Pistolenschüssen eine negative Wirkung, mit den anderen Einschüben (Gerüche, Lichtblitze, Bilder, Schwindelgefühle) konnte er keine Veränderung bewirken. Wenn das Ereignis nach dem Lernen auftrat, war die Wirkung größer.

In einer Reihe ähnlich aufgebauter Versuche wurden teils hemmende und teils fördernde Einflüsse gefunden. Daraus entstand die Hypothese, daß mäßige Gefühle förderlich, starke Gefühle dagegen störend sind.

Zu der Auswirkung von induziertem Erfolgs- oder Mißerfolgserleben liegt eine Reihe von Untersuchungen vor. SULLIVAN (1927) ließ Kinder sinnlose Silben lernen und gab ihnen entweder eine positive oder eine negative Rückmeldung. Durch das Mißerfolgserlebnis wurde die Zeit zum Auswendiglernen verlängert.

ZELLER (1950 b, 1951) veranlaßte seine Vpn sinnlose Silbenpaare bis zur völligen Beherrschung zu lernen. Die Kontrollgruppe erhielt bei einem Zwischentest (eine andere Aufgabe) eine neutrale Leistungsrückmeldung, die Experimentalgruppe ein äußerst negatives Feedback ("totale Versager"). Bei der anschließenden Retentionsprüfung schnitt die E-Gruppe viel schlechter ab als die K-Gruppe, sie konnten nur wenige Silben reproduzieren. Bei einem zweiten Zwischentest erhielten die Vpn der E-Gruppe eine positive Rückmeldung (Aufhebung des Mißerfolgs). Jetzt zeigten sie wieder bessere Behaltensleistungen, die fast das Niveau der K-Gruppe erreichten. ZELLER sah seine Er-

gebnisse im Einklang mit der Verdrängungstheorie von FREUD.

HOLMES (1972) erzielte auch durch eine positive Rückmeldung eine Verschlechterung der Behaltensleistung. Er zweifelte deshalb die Verdrängungstheorie an und favorisierte die Interferenztheorie. Auch BOCK (1978 a) meint, daß der Befund von HOLMES "gegen eine pa Verdrängungstheorie und für eine Art Interferenz-Theorie (spricht): Wenn Vpn emotionalisierenden Erfahrungen ausgesetzt werden, nachdem sie die zu behaltenden Items erhalten haben, stört dies rückwirkend das Behalten." (S.9)

ZÖHRER (1970) konnte zeigen, daß emotionalisierende Inhalte nicht nur eine retroaktive Hemmung hervorrufen können, sondern auch eine proaktive Hemmung ausüben. Photographien von Mordopfern, die den Vpn vor der Lernaufgabe gezeigt wurden, beeinträchtigten das nachfolgende Behalten von Paarassoziationen (Bild-Figur Kombinationen) stärker als neutrale Bilder.

Auf die Versuche zur Erinnerung an unterbrochene, abgebrochene (ZEIGARNIK 1927, ROSENZWEIG 1938) oder nichtgelöste (DUTTA & KANUNGO 1975) Aufgaben, von denen vermutet wird, daß auch der Gefühlsfaktor dabei eine Rolle spielt, wird in dieser Arbeit nicht eingegangen. Sie werden ausführlich bei RAPAPORT (1977) und DUTTA & KANUNGO (1975) behandelt.

BOCK (1978 a) zieht aus den Untersuchungen folgende Schlußfolgerung: Angenehme und unangenehme Erfahrungen "verschlechtern das Behalten, wenn sie vor oder nach den Ereignissen auftreten, deren Behalten überprüft wird (pro- oder retroaktive Hemmung). Treten sie dagegen simultan mit den zu behaltenden Ereignissen auf oder sind sie mit diesen identisch, verbessern sie das Behalten." (S. 10). Dies ist eine zu stark vereinfachte Sichtweise. Die Daten der Untersuchung von ZÖHRER (1970) sprechen dagegen. Grauererregende Bilder und neutrale Wörter wur-

den den Vpn gleichzeitig dargeboten. Die Behaltensleistung bei den Wörtern war schlechter verglichen mit einer Kontrollgruppe, die neutrale Bilder und Wörter sah.

Betrachtet man zusammenfassend die Versuche, die vor allem zu Beginn unseres Jahrhunderts durchgeführt wurden, so fällt besonders auf, daß sie teilweise große methodologische Mängel aufweisen. Meist fehlte die Kontrolle von wichtigen Störvariablen, die Anzahl der Vpn war häufig niedrig. Oft fand keine adäquate statistische Behandlung der Daten statt. Vor allem aber existierte keine explizit formulierte Theorie. All diese Faktoren führten zwangsläufig zu vielfach unterschiedlichen, sich widersprechenden Ergebnissen.

RAPAPORT (1950/1977, S. 137,138) zieht nach seinem Überblick über die experimentellen und klinischen Arbeiten folgendes Resümee:

Uns hat das Versuchsmaterial gezeigt, daß 'Gefühlsfaktoren' das Gedächtnis nachweisbar beeinflussen. Dieser nachweisbare Einfluß beschränkt sich nicht auf die Menge des gelernten Materials, die reproduzierbar ist, sondern schließt auch das Wiedererkennen und das Wiedererlernen ein; er berührt sowohl die Reproduktion und das Wiedererkennen von persönlichen Erlebnissen als auch die Reaktionszeit von Assoziationen und Sich-ins-Gedächtnis-Rufen, die Abfolge des Auftauchens von Material beim Sich-Erinnern und bei der freien Aufzählung...

Wie der Einfluß im einzelnen aussieht, ist bis heute noch nicht geklärt. Daß die Gefühlsintensität vermutlich wichtiger ist als die Gefühlsqualität betont auch RAPAPORT, ebenso wie BOCK (1978 a) und DUTTA & KANUNGO (1975). Nach RAPAPORT wurden konsistentere Ergebnisse erzielt, wenn die Bestimmung der Affektbetonung individuell vorgenommen wurde.

1.4. Theoretische Positionen

1.4.1. Qualitätsorientierte Positionen

DUTTA & KANUNGO (1975) gliedern die theoretischen Betrachtungen, die sich auf die Qualität der Affekte beziehen, in

fünf Hauptpositionen:

1. Hedonistische Position

Sie postuliert allgemein das Lustprinzip. Jeder Organismus sucht das Angenehme und vermeidet das Unangenehme. Die Verdrängungstheorie von FREUD wird als Spezialfall hier subsumiert.

2. Spannungssystem-Position (LEWIN 1935, ZEIGARNIK 1938 u.a.)

Sie sieht das selektive Behalten in Zusammenhang mit nicht entladenen Spannungssystemen.

3. Persönlichkeitstyp-oder differentielle Position

Sie erklärt das selektive Erinnern angenehmer und unangenehmer Erfahrungen in Begriffen unterschiedlicher Persönlichkeitsorientierungen oder -dispositionen, die die Individuen dazu bringen entweder die 'Ich-bedrohenden' oder die 'Ich-erhöhenden' Aspekte ihrer Erfahrungen zu betonen. [Gedächtnisoptimismus (KOWALEWSKI 1904); 'Ich-Stärke' (ALPER 1948); 'Bedürfnis nach Leistung' (ATKINSON 1953); Introversion - Extraversion (ERIKSON 1954)]

4. Kontextuelle Position

Der Kontext entscheidet, welches Material in Erinnerung bleibt. (BARRETT 1938: 'emotional set')* Die Übereinstimmung mit dem vorhandenen 'geistigen Bezugsrahmen' der Individuen ist die entscheidende Variable. (EDWARDS 1942: 'frame of reference')

5. Position der 'retroaktiven Hemmung'

Sie vertritt die Ansicht, daß U-Material ungeeignete und sich widerstreitende Reaktionen hervorruft, was zur Interferenz führt. (STAGNER 1931, D'ZURILLA 1965)

Diese fünf theoretischen Positionen vermuten, daß die Variation in der affektiven Qualität der Erfahrungen verantwortlich ist für das selektive Behalten. Jede schlägt aber einen anderen Mechanismus vor, um die Beziehung zwischen affektiver Qualität der Erfahrungen und ihrem Behalten zu erklären.

* Vpn, die unmittelbar vor einer Prüfung standen, erinnerten mehr U-Material.

1.4.2. Intensitätsorientierte Positionen

Zu den Forschern, die die Intensität der Affekte als ausschlaggebenden Faktor ansehen, zählen DUTTA & KANUNGO (1975), LIPMANN (1911), JERSILD (1931), MENZIES (1935) und viele andere.

LIPMANN nahm an, daß durch die Intensität der Gefühle die Intensität der Gedächtnisspuren bestimmt würde. Intensivere Spuren verschwinden demnach weniger rasch und werden leichter aktualisiert.

Eine ähnliche Ansicht kommt in der perseverativen Konsolidierungstheorie (WALKER 1958 und WALKER & TARTE 1963) zum Ausdruck, die 'Arousal'* zum Gedächtnis in Beziehung setzt. Eine höhere Erregung führt zu einer stärkeren Konsolidierung und dadurch zu bleibenderen Gedächtnisspuren.

Die 'perceived intensity of affect' Hypothese von DUTTA & KANUNGO (1975) sieht folgendermaßen aus: Die von einem Individuum wahrgenommene Gefühlsintensität bestimmt unabhängig von der Qualität das selektive Lernen und Behalten.** Die Gefühlsintensität hängt ab von der motivationalen Involviertheit des Individuums, seinen Erwartungshaltungen und Persönlichkeitscharakteristika und vom Kontext. DUTTA & KANUNGO formulieren eine positiv lineare Beziehung zwischen Gefühlsintensität und Behaltensleistung. Je intensiver das wahrgenommene Gefühl bei einer Erfahrung ist, desto besser wird sie behalten.

Eine 'positiv linear' gedachte Beziehung zwischen Intensität und Gedächtnis ist häufig anzutreffen. Es taucht aber auch

* BOCK (1978 a) spricht sich für eine neuropsychologische Konsolidierungstheorie aus. Emotionsstärke wird mit dem psychophysischen Arousal-Niveau gleichgesetzt.

** DUTTA & KANUNGO erfassen die individuelle Gefühlsstärke mit Hilfe von Ratingskalen. In ihrer Hypothese geht es ihnen um die von der Vp wahrgenommene Intensität. Sie äußern am Ende ihres Buches die Vermutung, daß vielleicht die physiologischen Korrelate der Erregung ein besseres Maß liefern. Die vorliegende experimentelle Arbeit soll ein Schritt zur Lösung dieses Problems sein.

die Vermutung auf, daß extreme Gefühle schädlich sein können (umgekehrte U-Funktion). Eventuell ist hier die Aufmerksamkeit nur noch auf die Gefühle gerichtet und nicht mehr auf die auslösende Ursache.

Über die genaue Wirkungsweise von Emotionen auf das Gedächtnis herrscht bis heute noch keine Klarheit. Ob die Wirkungen direkter oder indirekter Art sind, ist noch offen. Für eine eher direkte Einwirkung spricht die Ansicht, daß die Erregung Energie für die Konsolidierung zur Verfügung stellt. Einen indirekten Zusammenhang vermutet z.B. GORDON (1905), die als Vermittlungsglied zwischen Gefühlen und Gedächtnis die Aufmerksamkeit beschreibt. Die von Gefühlen beeinflussbare Aufmerksamkeit ist für die Gedächtnisleistung verantwortlich.

Schlußbemerkung

Bis heute fehlt eine Gedächtnistheorie, in der der Einfluß von Emotionen explizit dargestellt wird.

Das Buch von RAPAPORT (1950/1977) stellt einen Versuch dar, zur Entwicklung einer Gedächtnistheorie beizutragen, "in der Gedächtnisphänomene von der emotional-affektiven Organisation abhängig sind." (S.171), die also auf emotionalen Faktoren beruht. Er entwickelt letztlich aber selbst keine Theorie, sondern stellt nur die Forderung nach einer solchen Theorie auf. RAPAPORT vermutet, daß es "eine ganze Hierarchie von emotionalen Faktoren gibt, angefangen von den echten und tief verwurzelten bis hin zu den intellektualisierten und konventionalisierten." (S.321) Notwendig wären seiner Meinung eine genaue Untersuchung einer solchen Hierarchie und eindeutige Definitionen der verschiedenen Begriffe. Weiterhin fordert er, daß die Beziehung zwischen 'empfundene[n] Gefühl' und physiologischen Vorgängen geklärt werden muß.

1971 (im Vorwort zur 5.Auflage des Buches von RAPAPORT 1977) schreibt LESTER LUBORSKY, daß die von RAPAPORT geforderte neue Gedächtnistheorie, in der die Rolle emotionaler Einflüsse unerläßlich ist, bis heute noch nicht formuliert wurde und meint: "Gefühle und Motivation werden in Theorien der Erinnerungsvorgänge immer noch zu wenig betont". (S.12)

2. Die emotionale Tönung von Wörtern

Die experimentelle Untersuchung in dieser Arbeit soll den Einfluß des emotionalen Erregungspotentials von Wörtern auf die Gedächtnisleistung überprüfen. Als Stimuli dienen Wörter. Es erscheint daher angebracht, an dieser Stelle einigen Fragen hinsichtlich der Beziehung zwischen Gefühlen und Wörtern nachzugehen.

Daß Wörter Emotionen auslösen können, darauf gibt es viele Hinweise. Man denke dabei nur an die Reaktionen, die durch Tabuwörter, Reizwörter oder Schimpfwörter hervorgerufen werden. Gerade Beleidigungen haben in manchen Fällen sogar ein gerichtliches Nachspiel.

Vermutlich werden durch mehrere zusammenhängende Wörter, Sätze oder Texte eher und stärkere Emotionen ausgelöst als durch einzelne Wörter. Als Beispiele dafür, wie mit Worten Emotionen geweckt werden, seien aufreißerische Zeitungsschlagzeilen, Drohungen, Wahlkampfreden, Märchen, Gruselgeschichten, Liebes- oder Kriminalromane genannt.

Doch, wie schon erwähnt, kann auch das einzelne Wort Gefühle hervorrufen; vor allem gilt das für die Inhaltswörter unserer Sprache. FRÖBES (1929, S.287) schreibt: "Selbst bloße Worte können Gefühle wecken, wie Waldesrauschen, Giftmord, Freiheit, hochherzig, abscheulich. Der Reiz einer Erzählung beruht zum großen Teil auf solchen durch Worte geweckten Gefühlen."

So wie Wörter Emotionen auslösen können, so lassen sich auch über Worte Emotionen abreagieren (schimpfen, streiten, usw.). Dieser zweite Aspekt soll in dieser Arbeit aber ausgespart bleiben.

Die Wörter symbolisieren Gegenstände, Vorgänge, Ereignisse und Zusammenhänge. Ein Wort ruft nicht nur eine isolierte Vorstellung des Bezeichneten hervor, sondern es weckt gleichzeitig weitere Empfindungen, die damit verbunden sind: Erlebnisse, Erfahrungen, Gefühle, Einstellungen, die mit dem

Bezeichneten zusammenhängen. So äußert W.SMITH (1922/1970, S.27): "When a stimulus-word is applied, a host of associated images of one kind and another is called up in the subjects mind of a nature dependent on his past experience." ERDMANN (1900, zit. nach FUCHS 1975 a, S.58) unterscheidet am Wort den begrifflichen Inhalt, den Nebensinn und den "Gefühlswert oder Stimmungsgehalt" (= "alle reaktiven Gefühle und Stimmungen, die es erzeugt"). Die Trennung in den "sachlichen Inhalt eines Begriffs" und in das, "was mit an klingt", wenn man ein Wort hört oder spricht, wird durch die Bezeichnungen 'denotative und konnotative Bedeutung' ausgedrückt. (HÖRMANN 1977, S.111)

HÖRMANN (1977) weist darauf hin, daß der rationale, sachliche (denotative) Aspekt der Sprache des Hauptthema aller traditionellen Sprachtheorien bildete, angefangen bei ARISTOTELES bis WUNDT.

Der Sprachphilosoph STENZEL (1934, zit. nach BOTTENBERG 1972, S.157) äußert den Gedanken einer emotionalen Grundlage sprachlicher Bedeutung.

Das sprachpsychologische Modell von OSGOOD beschäftigt sich ebenfalls mit diesem Ansatz. In seiner Mediationstheorie der Bedeutung sieht OSGOOD das Vermittlungsglied zur Bedeutung eines Wortes im emotionalen Bereich (im Gegensatz zu BOUSFIELD, der es im Bereich des Verbalen sieht). HÖRMANN (1977, S.113) beschreibt OSGOODs Ansicht folgendermaßen: "Zeichen erlangen Bedeutung nicht nur durch ihre Verbundenheit mit den mit ihnen assoziierten Wörtern, sondern auch und vor allem durch Verbindung mit nicht primär verbalen, sondern emotionalen, einstellungsmäßigen Beständen." OSGOOD und Mitarbeiter (1957) entwickelten ein Meßverfahren zur Erfassung der Mediatisierungsprozesse, das 'Semantische Differential'. Mit Hilfe des Semantischen Differentials werden die grundlegenden Dimensionen der Bedeutung erfaßt. Auf bipolaren, meist 7-stufigen Adjektivskalen beurteilt die Vp Begriffe. In einer Reihe von Faktorenanalysen erhielt OSGOOD immer wieder folgende drei Hauptfaktoren, die in der Regel mehr

als 2/3 der gemeinsamen Varianz erklären (FUCHS 1975 b, S.85):

1) Bewertung (Evaluation), 2) Stärke (Potency), 3) Aktivität (Activity). Sie sind die grundlegenden Dimensionen der konnotativen Bedeutung und entsprechen nach OSGOODs Meinung generellen Gefühlsdimensionen. Er weist auf die Übereinstimmung seiner drei Faktoren mit den drei Grunddimensionen des Gefühls nach WUNDT hin (Lust - Unlust, Spannung - Lösung, Erregung - Beruhigung).

Den Aspekt des emotionalen Erregungspotentials, dem in dieser Arbeit nachgegangen werden soll, kann man unter den konnotativen Aspekt subsummieren. Er ist allerdings enger gefaßt. Bevor eine Definition von meiner Seite erfolgt, soll betrachtet werden, wo man sich in der Wissenschaft mit der emotionalen Tönung von Wörtern befaßt, welche Bezeichnungen und welche operationalen Definitionen verwendet werden.

Neben den sprachpsychologisch orientierten Untersuchungen, die mit Hilfe der Assoziationsmethode und Korrelationsberechnungen in bezug auf andere Wortcharakteristika versuchen, mehr Aufschluß über den Aspekt der affektiven Tönung von Wörtern zu erhalten, werden 'emotionale' Wörter verwendet, um andere Phänomene zu untersuchen, wie z.B. den Einfluß von Emotionen auf das Gedächtnis oder den auf die Wahrnehmung. Zu nennen sind hier Wahrnehmungsuntersuchungen, die unter der Bezeichnung 'perceptual defense' bekannt geworden sind. Dabei werden nicht die Wörter selber analysiert, sondern man benützt sie als Mittel zum Zweck. Ein weiteres Gebiet, auf dem neben anderen Stimuli auch Wörter herangezogen werden, befaßt sich mit der Beziehung zwischen physiologischen Reaktionen und subjektiven Einschätzungen der Empfindungen. Auch in der Werbepsychologie verwendet man die emotionale Bedeutung von Wörtern, um z.B. Markennamen zu konstruieren. (Beispiel: aus 'ideal' wird 'Elida' oder 'Delial')

Ein anderer Anwendungsbereich darf natürlich nicht vergessen werden: die psychoanalytischen Assoziationsversuche vor allem zu Beginn unseres Jahrhunderts. Bei diesen diagnostischen Assoziationsversuchen, besonders ist hier JUNG

(1904) zu erwähnen, wurde der Aspekt ausgenutzt, daß Wörter auf gefühlsbeladene 'Komplexe' verweisen können. Dem Probanden wurde eine Liste von 100 Wörtern zweimal hintereinander vorgelesen, und er sollte auf jedes Wort mit dem ersten ihm einfallenden Wort reagieren. Beim 2.Durchgang sollte er versuchen, mit den gleichen Wörtern zu antworten wie beim 1.Mal. Als Hinweis, daß ein Wort einen 'Komplex' ansprach, galten folgende Indikatoren:

- 1) verlängerte Reaktionszeit
- 2) Blockierung (keine Reaktion)
- 3) Wiederholung des Reizwortes
- 4) sinnlose Klangassoziation, andere ungewöhnliche Assoziation
- 5) beim 2.Durchgang eine falsche Reproduktion*.

PETERSON & JUNG (1907), W.SMITH (1922/1970) u.a. bezogen Hautwiderstandsmessungen beim Assoziationstest mit ein und fanden, daß bei emotional belasteten Wörtern größere psychogalvanische Reaktionen auftreten.

2.1. Bezeichnungen in der Literatur

Im folgenden sollen verschiedene gebräuchliche Bezeichnungen angeführt werden, die man in der Literatur vorfindet. Man kann die Bezeichnungen in drei Gruppen einteilen, 1. in allgemeine, deren Gegensatz 'neutral' oder 'indifferent' lautet, 2. in solche, die die Qualität ansprechen und 3. in jene, die den Intensitätsaspekt vertreten.

2.1.1. Allgemeine Bezeichnungen

Hier tauchen die Begriffe 'emotional' oder 'affektiv' auf, entweder alleine oder in Verbindung mit 'getönt', 'beladen',

* Letzteres deutet auf eine Störung der Behaltensleistung bei stark gefühlsbetonten Stimuli hin. R.M.JONES (1958, zit. nach CRAMER 1968) wies nach, daß die Behaltensleistung nicht nur bzgl. der Reaktion vermindert war, sondern auch bzgl. des Stimulus, wenn man der Vp ihre Assoziation vorgab und sie mit dem Stimuluswort reagieren sollte.

'behaftet', 'gefärbt', 'besetzt', 'bedeutsam'. Daneben gibt es diese Zusammensetzungen auch mit 'gefühls-' (z.B. gefühlsbetont) oder 'emotions-' (z.B. emotionsgeladen).

Diese Bezeichnungen werden entweder als Adjektive vorangestellt oder es erfolgt die Klassifizierung der Wörter durch Nachstellen von 'mit...', also: Wörter mit emotionalem Wert, Ton, Inhalt, Gehalt oder mit emotionaler Konnotation, Resonanz, Bedeutung.

Beispiele:

emotionale Wörter (EMRICH & HEINEMANN 1966)
emotional words (KELLEY 1968, MANNING & JULIAN 1975, SINHA & SINHA 1977 u.v.a.)
affective words (CRAMER 1968, RHONE 1975, GUERNETTE 1975)
mots affectogènes (MOSCU 1971, 1973)
emotional besetzte Wörter (JANKE in SCHÖNPFLUG 1969)
emotional bedeutsame " (EMRICH & HEINEMANN 1966)
emotional-loaded words (MOSCU 1971)
emotionally-charged words (DiCAPRIO 1969)
emotionally toned words (YARMEY 1966)
affectively laden words (KAYTON & KROH 1975)
affectively-toned words (ROUTH & ELLIS 1975)
emotionsgeladene Wörter (BOTTENBERG 1972)
gefühlsbeladen, gefühlsbehaftet, -betont (EMRICH & HEINEMANN)
affect-laden words (MOSCU 1973) 1966
words of emotional value (SYZ 1926/27 a, WECHSLER & JONES 1928)
affective value (JONES & WECHSLER 1928)
emotional content (MANNING & JULIAN 1975)
affective content (BASSOS 1973)
emotional connotation (HUTT & ANDERSON 1967)
affective connotation (McNULTY & ISNOR 1967)
emotive tone (YARMEY 1966)
affective tone (W.SMITH 1922/1970, McGEOCH & IRION 1952)
affective resonance (MOSCU 1971)
affective intensity (GUERNETTE 1975)
emotionale, affektive Bedeutung (FUCHS 1975)
affective meaningfulness (AUGUST & FELKER 1977)
emotionale Ladung, Potenz (REYKOWSKI 1973)
emotionaler Bedeutungsgehalt (FUCHS 1975)

Es gibt aber auch Bezeichnungen, bei denen das Adjektiv 'emotional', 'affektiv' oder 'gefühlsbetont' fehlt.

Beispiele:

Wörter mit hoher Valenz (SCHÖNPFLUG 1966)
Wörter mit hoher Ich-Bezogenheit (SCHÖNPFLUG & BEIKE 1964)
(persönlich) bedeutsame Wörter (SCHÖNPFLUG & BEIKE 1964)
personally relevant Words (DiCAPRIO 1969)
critical words (WECHSLER & JONES 1928, JONES & WECHSLER 1928)

2.1.2. Qualitativ orientierte Bezeichnungen

Hier werden die emotionalen Wörter hinsichtlich ihrer Qualität spezifiziert. Häufig findet sich nur die grobe Klassifizierung in 'positiv' und 'negativ'. Früher verwendete man dafür die Bezeichnungen 'lustbetonte' und 'unlustbetonte Wörter'. Am häufigsten benutzt wird jedoch die Klassifizierung 'angenehme' und 'unangenehme Wörter', im englischen Sprachraum 'pleasant - unpleasant words' oder 'good - bad words'.

Beispiele:

positive - negative words (RHONE 1975, GUERNETTE 1975)

pleasant - unpleasant words (ANISFIELD 1970 u.v.a.)

good - bad words (AMSTER 1964)

liked - disliked words (AUGUST & FELKER 1977)

words of positive tone - negative tone (W.SMITH 1922/1970)

pleasant - unpleasant connotation (THOMSON 1930)

affectively positive - negative meaningfulness

(RYCHLAK et al. 1974)

Zu den qualitativen Bezeichnungen gehören auch solche, die die angesprochenen Gefühle nur auf bestimmte einschränken, wie es z.B. bei den 'upsetting words' von MANNING & MELCHIORI (1974) und MANNING & GOLDSTEIN (1976) der Fall ist. Es handelt sich dabei um Wörter, die die Vp aus der Fassung bringen und bei ihr negative Emotionen wie Angst auslösen sollen. Die Definition von 'upsetting words' lautet: "An upsetting word is a word with a connotation which causes the individual who is reading about, thinking about, or using the particular word in speech to experience anxiety and thus suffer some kind of discomfort." (MANNING & GOLDSTEIN 1976, S.243) Ursprünglich zählt MANNING die Tabuwörter ('taboo words') mit zu den 'upsetting words', später behält sie die Bezeichnung 'upsetting' nur für die 'non-taboo emotional words' bei und behandelt die Tabuwörter als eigene Kategorie. Tabuwörter sind bei ihr Wörter aus dem sexuellen ('sex') und dem Toilettenbereich ('bathroom').*

* Letzterer Bereich ist in den USA stark tabuisiert. Wenn man dort zur Toilette will, fragt man "where can I wash my hands".

Neben der Spezifizierung bzgl. der emotionalen Reaktion, die das Wort auslöst, wie z.B. 'angstauslösende Wörter' oder 'peinliche Wörter', werden Spezifizierungen nach dem emotional besetzten Bereich vorgenommen, dem das Wort inhaltlich angehört. Diese Wörter rufen nicht unbedingt die Emotion hervor, für die sie stehen. 'Aggressive Wörter' (= Wörter mit aggressivem Inhalt) können z.B. auch Angst auslösen.

Beispiele:

aggressive words (CAMP & ZIMET 1973, GORDON 1976)
hostile words (BODIN & GEER 1965, zit. nach CRAMER 1968)
death words (den Tod betreffende Wörter) (TEMPLER 1971)
sex words (MANNING & GOLDSTEIN 1976)

2.1.3. An der Intensität orientierte Bezeichnungen

Dazu gehören jene, die den Begriff 'Arousal' verwenden. Meist zeigt sich nur eine Dichotomie; die Wörter zerfallen in stark vs. schwach erregende. Selten werden mehrere Grade spezifiziert.

Beispiele:

high arousal words - low arousal words (WALKER & TARTE 1963)
most arousing word, medium arousal words, low arousal words (POWERS 1974)

Man könnte hier auch die Bezeichnungen 'hohe' vs. 'niedrige Ich-Bezogenheit' aufführen, da durch 'hoch' und 'niedrig' eine Abstufung in der Stärke angedeutet ist.

Eine weitere Bezeichnung ist die 'emotionality of words', die sich nur auf die Intensität unabhängig von der Art der Qualität bezieht. Bei BROWN & URE (1969) als E abgekürzt, bei NOBLE (1958) als e.

An den Bezeichnungen, die in der Literatur auftauchen, läßt sich ablesen, daß neben der globalen Kennzeichnung 'emotional' vs. 'nicht emotional' oder 'neutral' sowohl der Aspekt der Qualität als auch der Aspekt der Intensität der Emotionen für sich allein betrachtet wird.

2.2. Operationalisierungen

Wie wird nun die Emotionalität eines Wortes bestimmt? Die Gliederung der folgenden Besprechung der verschiedenen Bestimmungs- und Meßmethoden orientiert sich an der üblichen Einteilung in die drei Meßebenen zur Erfassung von Emotionen: 1. die subjektive, 2. die physiologische, 3. die Verhaltensebene.

2.2.1. Subjektive Einschätzungen (SE)

Die subjektive Beurteilung ist die am häufigsten verwendete Methode. Mit ihr wird alles erfaßt: die allgemeine Emotionalität eines Wortes (das Wort ist emotional oder neutral), die emotionale Qualität eines Wortes (das Wort ist angenehm oder unangenehm, oder spezifischer: das Wort ruft Angst hervor; dieses Wort ist aggressiv usw.) und schließlich die emotionale Intensität eines Wortes (das Wort ist stark erregend, schwach erregend).

a) durch den V1

Die Beurteilung wird im einfachsten Fall vom Experimentator vorgenommen. Er bestimmt, welches Wort emotional und welches neutral ist, welche Wörter angenehm und welche unangenehm sind. Diese Art der Bestimmung wurde hauptsächlich in früheren Untersuchungen angewendet (TAIT 1913, TOLMAN 1917, LYNCH 1932; zit. nach RAPAPORT 1977). Das Verfahren ist nicht unbedenklich, weil nicht von vornherein angenommen werden kann, daß die Empfindungen von V1 und Vp deckungsgleich sind. Das Urteil wird nur von einer Person vorgenommen und auf alle anderen Vpn übertragen. Der V1 kann dabei bestimmte Wörter als emotional erleben, die für seine Vpn neutral sind. In dieser Vorgehensweise ist die Annahme enthalten, daß alle Menschen in gleicher Weise bestimmten Wörtern eine bestimmte Emotionalität zuweisen würden.

Neben diesem Absolutheitsanspruch existiert aber auch die Sichtweise, daß die Emotionalität eines Wortes von der Vp

abhängt und je nach Vp andere Wörter emotional sind. Aber auch in diesem Falle traut sich der V1 zu, die Emotionalität der Wörter zu beurteilen. So sieht MOSCU (1971) für ihre Vpn, die Wettkampfsportler sind, solche Wörter als affektiv an, die deren sportliche Aktivität betreffen, vor allem, wenn die Vpn kurz vor dem Wettkampf stehen. Im Jahr 1973, als sie Schauspieler untersucht, verwendet sie als emotionale Wörter solche, die mit dem Auftritt auf der Bühne zu tun haben.

Die Neutralität eines Wortes zu beurteilen scheint weniger Schwierigkeiten zu bereiten als die Emotionalität, da in manchen Untersuchungen, in denen die Vpn die emotionalen Wörter bestimmten, die Auswahl der neutralen Wörter nur vom V1 getroffen wurde (MANNING & MELCHIORI 1974).

b) durch unabhängige Beurteiler

Die Affektivität eines Wortes wird durch einige unabhängige Beurteiler festgestellt. Der V1 nimmt also die Beurteilung nicht alleine vor, sondern befragt einige 'Experten'. So wurde z.B. die Aggressivität von Wörtern von fünf Mitgliedern einer Forschungsgruppe bestimmt (CAMP & ZIMET 1973). Die Wörter, bei denen die größte Übereinstimmung vorliegt, werden als emotional oder je nach Beurteilungskategorie als angenehm, unangenehm usw. angesehen, oder es erfolgt eine Mittelung der Werte über die Beurteiler. Die Emotionalität gilt auch hier wieder für alle Vpn.

c) durch die Vp

Die Vp oder die Vpn schätzen die Wörter selber ein. In den bisher angeführten Fällen, in denen entweder der V1 oder eine Gruppe von Fachleuten die Wörter beurteilten, galt dieses Urteil für alle Vpn. Hier kann man die Daten entweder normativ oder ipsativ auswerten. Bei der normativen Bestimmung werden pro Wort die Bewertungen aller Vpn zusammengefaßt und anhand des Gruppenurteils die Wörter klassifiziert. Das Wort erhält entweder den Durchschnittswert der Beurteilungen oder es werden nur die Wörter klassifiziert, bei de-

nen übereinstimmende Beurteilungen vorliegen. CHANEY & LAUER (1929, zit. nach RAPAPORT 1977) ließen 150 Vpn Wörter als A, U und N bewerten. Die Wörter mit der größten Übereinstimmung wurden ausgewählt.

Bei der ipsativen Methode gilt die Einschätzung der Vp nur für sie persönlich. Ein und dasselbe Wort kann für eine Vp emotional und für eine andere neutral sein. Die Urteile werden hier nicht zusammengefaßt, es bleibt bei Einzelurteilen. STAGNER (1933, zit. nach RAPAPORT 1977) geht bei seiner Untersuchung in dieser Weise vor.

RHONE (1975) unterscheidet die beiden Verfahren terminologisch folgendermaßen: 'group-rated words' vs. 'individually rated words' (Gruppenwerte vs. individuelle Werte).

Die subjektive Beurteilung durch die Vp zeigt sich nicht nur bei der direkten Befragung der Vp (es werden ihr Wörter vorgelegt, die sie einschätzen soll) sondern auch, wenn die Vp die Wörter selbst produzieren soll. Die Vp wird gebeten, emotionale Wörter zu nennen oder aufzuschreiben. THOMSON (1930) forderte ihre Vpn auf "to make a list of twenty words which had a very pleasant connotation for him and twenty a very unpleasant one" (S.464). MANNING et al. (1974, 1975, 1976) erhoben ebenfalls auf diese Weise ihre Wörter. Die Vpn sollten (1974) die drei beunruhigendsten ('upsetting') Wörter aus verschiedenen Bereichen notieren. (THOMSON verarbeitete ihre Daten ipsativ, MANNING et al. normativ, die häufigst genannten wurden verwendet.)

Ein weiterer Sonderfall liegt vor, wenn der Vp eine große Anzahl Wörter vorgelegt werden und sie nicht jedes klassifizieren muß, sondern nur bestimmte Wörter (z.B. emotionale, neutrale, angenehme, unangenehme usw.) auswählen soll.

Bei allen drei Verfahren äußert sich aber letztendlich das subjektive Urteil, deshalb werden sie unter der Bezeichnung 'Subjektive Einschätzung' aufgeführt, obwohl es sich strenggenommen bei den letzten beiden (Produktionsmethode und Auswahlmethode) nicht um reine Einschätzungen handelt.

2.2.1.1. Skalen

Die unterschiedlichen Skalen, mit denen gearbeitet wird, sollen im folgenden dargestellt werden.

Die grobe Kategorisierung (emotional - neutral und angenehm - unangenehm) wird vom V1, von unabhängigen Beurteilern oder von Vpn vorgenommen. Die differenzierteren Beurteilungen stammen von mehreren Beurteilern oder von Vpn.

1. Allgemeine Beurteilung nach 'emotional' - 'neutral'

Die Kategorien 'emotional' und 'neutral' werden vorgegeben, und die Wörter sollen danach beurteilt werden. Eine geringfügige Variation: Die Vp soll die Wörter anstreichen, die bei ihr emotionale Reaktionen hervorrufen. So forderte SYZ (1926/1927 a) seine 64 Vpn auf, von 50 Wörtern die zu kennzeichnen, die bei der Darbietung emotionale Reaktionen bewirkt haben.

2. Einfache Beurteilung der emotionalen Qualität

Die Wörter sollen danach klassifiziert werden, ob sie als angenehm (lustbetont, gut), unangenehm (unlustbetont, schlecht) oder neutral (indifferent) empfunden werden. LANIER (1941, zit. nach WOODWORTH & SCHLOSBERG 1965) benutzte eine zusätzliche Kategorie für gemischte (mixed) Gefühle. Hier ordnete er Wörter ein, die einen affektiven Konflikt ausdrücken.

Andere Kategorien wie z.B. 'aggressiv', 'angsterzeugend', 'peinlich', 'beunruhigend' usw. zählen hier mit dazu.

3. Differenziertere Beurteilung der Qualität

Es gibt verschiedene Ratingskalen (5-, 7- oder 9-stufige), die teilweise sehr ähnlich sind. Verwendet werden hauptsächlich zwei bipolare Adjektivskalen mit den Polen 'pleasant' und 'unpleasant' oder 'good' und 'bad'. Neben der Anzahl der Stufen und der Bezeichnung der Skala zeigen sich nur Unterschiede bei der Zuordnung der Zahlen zu der positiven und negativen Seite. Bei einigen steht die 1 für 'angenehm', bei anderen für 'unangenehm'. Oft wird nur eine Skala herangezogen, manchmal aber auch mehrere (im Extremfall die evaluativen Skalen des semantischen Differentials). Die 7-

Punkte Skala (pleasant - unpleasant) wird am häufigsten benutzt (u.a. auch von McNULTY & ISNOR 1967).

Beispiele für einzelne Skalen:

DYSINGER (1973, zit. nach WOODWORTH & SCHLOSBERG 1965): 5-Punkte Skala mit den Stufen: very pleasant, pleasant, indifferent, unpleasant, very unpleasant.

AMSTER (1964): 9-Punkte Skala mit den Polen 'unpleasant' (1), 'pleasant' (9), in der Mitte 'neutral' (5).

SILVERSTEIN & DIENSTBIER (1968): 7-Punkte Skala 'Pleasantness' (Pl);

1	2	3	4	5	6	7
very	moderately	slightly	neutral	slightly	moderately	very
unpleasant						pleasant

BROWN & URE (1969) verwendeten die gleiche Skala wie SILVERSTEIN & DIENSTBIER; zusätzlich zu der 'Pleasantness'-Skala (P) benutzten sie eine 'Goodness'-Skala (G) mit den Polen 'bad' (1) und 'good' (7). Jede Skala wurde für sich genommen.

Beispiele für kombinierte Skalen:

FRINCKE (1968) faßte beide Werte der 7-Punkte Skalen G und P zusammen.

DUTTA & KANUNGO (1975) ließen ihre Vpn Adjektive auf 9-Punkte Skalen einstufen, mit den Polen 'good - bad', 'unpleasant - pleasant'. Bei einigen Versuchen verwendeten sie zusätzlich Skalen mit den Polen: 'favorable - unfavorable', 'attractive - repulsive', 'acceptable - unacceptable'. Beispiel ihrer Skala:

-4	-3	-2	-1	0	1	2	3	4
extreme	strong	moderate	mild	neutral	mild	moderate	strong	extreme
bad								good

Zur Anwendung kommen auch die bipolaren Adjektivskalen des Semantischen Differentials, die hoch auf dem evaluativen Faktor laden.

ANISFELD & LAMBERT (1966) und GUERNETTE (1975) wählten Wörter nach ihrer Ausstattung auf der evaluativen Dimension (bei HEISE 1965) aus.

4. Beurteilung der Intensität

Die Wörter werden entweder auf mehrstufigen Skalen (4 bis 7 Stufen) bzgl. der Stärke der ausgelösten Emotion skaliert, oder die Vp verteilt die Wörter, jeweils gleich viele (forced choice), auf die einzelnen vorgegebenen Intensitätsstufen (WECHSLER 1925).

Beispiele für Intensitätsskalierung:

WELLS & FORBES (1911, zit. nach WECHSLER 1925) gaben ihren beiden Vpn eine Skala vor mit den 4 Graden A: strongly emotional, B: rather emotional, C: rather unemotional, D: particularly devoid of any reaction.

SMITH & HARLESTON (1966) benützten eine 6-Punkte Skala 'of emotionality' (von 'extremely emotional' bis 'extremely unemotional').

PAIVIO (1968): 7-Punkte Skala E (emotionality), mit den Polen 'emotional' - 'unemotional'.

BROWN & URE (1969) maßen die Intensität mit ihrer 7-Punkte Skala (E), mit den Polen 'neutral' (1) und 'emotional' (7).

MANNING et al. (1974, 1975) verwendeten eine 5-Punkte Skala, auf der die Vpn beurteilten sollten, wie stark 'upsetting' das Wort für sie ist, wobei 1 für 'highly upsetting' und 5 für 'not upsetting' steht.

Die Bestimmung der Intensität wird auch mit Hilfe der Qualitätsurteile vorgenommen.

Beispiele:

NOBLE (1958) errechnet die Stärke der Emotionalität (e) anhand der qualitativen Urteile von 100 Vpn, indem er die vier Kategorien (neutral = N, pleasant = P, unpleasant = U und mixed = M) unterschiedlich gewichtet. Die Gewichtung basiert auf der Beziehung zwischen EDR-Amplitude und Kategorienurteil.

Formel:

$$e = \frac{1(f_N) + 2(f_P) + 3(f_U) + 4(f_M)}{n}$$

wobei n = Anzahl der Vpn, die das Wort beurteilen

f = Häufigkeit der Antworten in jeder Kategorie.

Die e-Werte schwanken zwischen 1 (niedrigster) und 4 (höchster e-Wert).

BROWN & URE (1969) versuchten auch eine Bestimmung der Intensität anhand der P und G Skalen. Die Einschätzungen bekommen dabei den Wert ausgedrückt in der Entfernung vom neutralen Punkt in der Mitte unabhängig von der Richtung.

DUTTA & KANUNGO (1975) ermittelten ebenfalls aus ihren qualitativen Skalen (von -4 nach +4) die 'perceived intensity of affect' durch Weglassen der Vorzeichen.

Neben der Methode der Beurteilung der Gefühle anhand einer Ratingskala mit willkürlich festgesetzten Stärkegrade, gibt es die Methode der Rangbildung. Die Vp ordnet dabei die Wörter in eine Rangfolge vom stärksten zum schwächsten erregenden Wort. (z.B. McCURDY 1950).

TRAXEL & BECHER (1957), die die obigen Methoden als unangemessenerachten, finden das Vergleichsurteil adäquater. Sie ließen deshalb 10 Wortpaare vergleichen, aber jeweils nur die zwei Wörter eines Paares (einfaches Vergleichsurteil). Die Vp sollte "beurteilen, welches von beiden die stärkere oder schwächere 'gefühlsmäßige' Wirkung gehabt habe ... und daß im Zweifelsfall mit 'indifferent' bzw. 'gleich' geurteilt werden könne." (S.291)

Die Paarvergleichsmethode, vor allem wenn alle Wörter miteinander verglichen werden sollen, ist aber nur durchführbar, wenn nicht allzuvielen Items zur Beurteilung anstehen. POWERS (1974), der diese Methode benutzte, hatte insgesamt nur sieben Wörter.

5. Beurteilung der konnotativen Bedeutung (nach OSGOOD)*

Methode des Semantischen Differentials:

Auf mehreren bipolaren, meist 7-stufigen Adjektivskalen beurteilen die Vpn einzelne Begriffe. Durch die Einschätzungen kann die Ausstattung der Wörter auf den drei Dimensionen 1) Bewertung (Evaluation), 2) Stärke (Potency) und 3) Aktivität (Activity) bestimmt werden.

Beispiele für Skalen:

Faktor 1:	gut	-	schlecht
	schön	-	häßlich
Faktor 2:	stark	-	schwach
	schwer	-	leicht
Faktor 3:	aktiv	-	passiv
	schnell	-	langsam

Die Dimension der 'Evaluation' entspricht weitgehend der affektiven Qualität (A und U). Der Unterschied liegt lediglich darin, daß die A oder U Qualität nur anhand einer Skala bestimmt wird, wohingegen sich der Faktor 'Evaluation' aus mehreren Adjektivskalen zusammensetzt, wobei meist die Skala 'angenehm - unangenehm' mit enthalten ist.

Fraglich ist, wie die Dimensionen 'Potency' und 'Activity' zu der affektiven Intensität stehen. Sicher sind sie nicht identisch mit ihr. Bei der Intensität handelt es sich um die

* Näheres über die Methode kann z.B. bei OSGOOD et al. (1957/1967) nachgelesen werden.

Stärke der durch ein Wort ausgelösten Emotion. Ein Begriff kann als 'schwach' oder 'passiv' beurteilt werden und trotzdem eine intensive Emotion bei der Vp hervorrufen. 'Activity' darf nicht gleichgesetzt werden mit der Gefühlsintensität oder mit der Aktivierung der Vp, die durch ein Wort ausgelöst wird.

Vergleich verschiedener Skalen

a) qualitative Skalen

Die Skalen G (goodness) und P (pleasantness) erfassen ähnliches, die Urteile unterscheiden sich kaum. Sie korrelieren in der Untersuchung von BROWN & URE (1969, S.235, 650 Wörter) .958 miteinander.

Der Vergleich der Einschätzungen von 135 Wörtern auf der G-Skala (BROWN & URE 1969) mit den Werten dieser Wörter auf der Dimension 'Evaluation' des Semantischen Differentials (HEISE 1965) ergab eine Korrelation von .87. Die Korrelation mit der P-Skala betrug .83. Die Skalen G und P sind bei den evaluativen Skalen des SD enthalten, deshalb ist dieser Zusammenhang nicht erstaunlich. Allerdings wird mit den evaluativen Skalen des SD nicht genau das gleiche gemessen.

b) qualitative und quantitative Skalen

Es zeigte sich kein Zusammenhang zwischen den qualitativen Urteilen und den Intensitätseinschätzungen. Beide Skalen messen also zwei unterschiedliche Aspekte des emotionalen Gehalts von Wörtern. Die Korrelation der E-Skala mit der G- und P-Skala betrug -.07 bzw. -.058 (BROWN & URE 1969, S.234). Dies ist verständlich, da bei der G- und P-Skala der neutrale Punkt in der Mitte der Skala liegt, bei der E-Skala hingegen auf einem der beiden Pole.

c) quantitative Skalen

BROWN & URE (1969) verglichen die Intensitätswerte ihrer E-Skala mit den aus den G- und P-Skalen gewonnenen Intensitätswerten. Die Polaritätswerte korrelierten signifikant positiv mit den E-Werten. ($G_{Pol.}$ und E: $r = .753$; $P_{Pol.}$ und E: $r = .702$). Da keine vollkommene Übereinstimmung vorliegt,

erfassen die Skalen unterschiedliche Intensitätsaspekte. Die G und P Polaritätswerte spiegeln die stärkere Ausprägung des Angenehmen oder Unangenehmen wider, während die E-Skalen die Intensität der Emotionalität unabhängig von deren Qualität messen.

2.2.1.2. Reliabilität

BROWN & URE (1969) führten Reliabilitätsbestimmungen der Einschätzungen durch. 50 Wörter wurden zweimal von der Hälfte der Vpn auf G-, P- und E-Skalen beurteilt. Die Reliabilitätskoeffizienten betragen für die Skalen G .994, P .996 und E .982. Die Übereinstimmung mit den Werten der anderen Hälfte der Vpn ergab .979 für G, .987 für P und .952 für E. (S.234)

Weiter überprüften sie ihre ermittelten G-, P- und E-Werte mit den Daten anderer Autoren bei den Wörtern, die sich in beiden Untersuchungen befanden. So zeigte sich bei 136 Wörtern, die auch im semantischen Atlas (JENKINS et al. 1958) aufgeführt sind, eine Korrelation von .93 bzgl. G. Der Vergleich ihrer P-Einschätzungen mit denen (Pl) von SILVERSTEIN & DIENSTBIER (1968) erbrachte für 45 Wörter, die bei beiden Untersuchungen verwendet wurden, eine Korrelation von .97. Die Urteile über 58 Wörter von BROWN (1965) und BROWN & URE zeigen auch eine gute Übereinstimmung. Die Korrelationen betragen .93 für P und .89 für E. Mit den E-Werten von LEVINGER & CLARK (1961) (60 gemeinsame Wörter) ist die Korrelation .91 und bei 60 Wörtern, die auch SMITH & HARLESTON (1966) einschätzen ließen, ergab sich eine Korrelation von .92. Die hohen Korrelationskoeffizienten sind der Tatsache zu verdanken, daß die Wortlisten heterogen waren (also emotionale und neutrale Wörter enthielten). Bei der getrennten Betrachtung der emotionalen (E) oder neutralen (N) Wörter (homogene Liste) fallen die Korrelationskoeffizienten entsprechend niedriger aus. So erhielten BROWN & URE bei der Aufteilung der Wörter in E- und N-Listen folgende Korrelationen beim Vergleich mit den Werten der letzten drei Untersuchungen:

BROWN (1965) LEVINGER & CLARK SMITH & HARLESTON

N	.88 (28 Wörter)	.77 (30 Wörter)	.44 (40 Wörter)
E	.75 (30 Wörter)	.83 (30 Wörter)	.41 (20 Wörter)

TRAXEL & BECHER (1957) führten mit neun Vpn nach ca. drei Wochen eine Wiederholungsprüfung der Intensitätsbeurteilung mit der Methode des einfachen Paarvergleichs durch. Diese erbrachte einen Stabilitätskoeffizienten von .60.

SMITH & HARLESTON (1966) ließen 60 Wörter auf einer 6-Punkte Skala bzgl. der Emotionalität von vier Gruppen von Vpn einschätzen. Alle Korrelationen lagen über .95 (S.312).

Die guten Übereinstimmungen sind vermutlich auch darauf zurückzuführen, daß die Stichproben (der Beurteiler) relativ homogen waren (Studenten als Vpn).

OSGGOD et al. (1967) und FUCHS (1975 b) berichten von zufriedenstellenden Reliabilitätswerten des Semantischen Differentials.

Einen Vergleich zwischen der Einschätzung von Vpn und der Ansicht der V1 über die Emotionalität von Wörtern kann man bei den Daten von MANNING & JULIAN (1975, S.243) anstellen. Als neutrale Wörter wurden Wörter aus den Bereichen Früchte (fruit), Getränke (beverage) und Lesematerial (reading material) zusammengestellt. Hier stimmt die Einschätzung (auf einer 5-Punkte Skala mit den Polen 1 = highly upsetting und 5 = not upsetting) mit der Intention überein. Bei den 'highly upsetting' Wörtern (nach der Ansicht der V1 und der Vpn, die die Wörter produzierten) ergab sich nur teilweise eine Entsprechung mit der Einschätzung durch andere Vpn. Nur die Wörter aus dem Bereich Gewalt (violence) wurden ähnlich beurteilt; die Wörter aus den Bereichen Sexualität (sex) und Toilette (bathroom) wurden gegensätzlich eingeschätzt, als nicht 'upsetting'.

<u>neutrale Wörter</u>	SE	<u>emotionale Wörter</u>	SE
fruit	5.00	violence	2.59
reading material	4.98	sex	4.40
beverage	4.92	bathroom	4.60

2.2.1.3. Probleme der Validität

Folgende Fragen müssen hier gestellt werden:

- Inwieweit kann die Emotionalität von Wörtern bewußt wahrgenommen werden?
- Inwieweit ist die Vp bereit, das Wahrgenommene ehrlich mitzuteilen?
- Inwieweit unterliegt die Vp Trugwahrnehmungen, Täuschungen, Vorurteilen und anderen Einflußgrößen?

Das Urteil ist abhängig von der Introspektionsfähigkeit und der Bereitschaft sich unbefangen zu beobachten und aufrichtig darüber zu berichten.

Manche Autoren äußern die pessimistische Ansicht, daß sich Gefühle nicht beobachten lassen; sie würden verschwinden, sobald man die Beobachtung auf sie lenkt. Zumindest eine Beeinträchtigung der Gefühle durch die Beobachtung wird häufig angenommen. Für diese Forscher wäre die Methode der subjektiven Einschätzung nicht geeignet, um den emotionalen Gehalt von Wörtern festzustellen. JANKE (in SCHÖNPFLUG 1969) sieht die Emotionalität von Wörtern als zu gering an und meint deshalb, daß sie nicht introspektiv erfaßt werden kann.

Unterstellt man aber, daß der emotionale Gehalt von Wörtern durch Selbstbeobachtung wahrgenommen werden kann, so muß berücksichtigt werden, daß das abgegebene Urteil von verschiedenen Faktoren beeinflusst werden kann. Es ist durchaus möglich, daß die Vp unrichtige Angaben macht und nicht bereit ist, ihre wahren Empfindungen zu offenbaren. Sozialpsychologische Faktoren können hier eine Rolle spielen. Die befragte Person möchte in einem günstigen Licht erscheinen, sie orientiert sich an bestehenden sozialen Normen und Konventionen.

Fehlerhafte Aussagen können auch durch Vorurteile und Selbsttäuschungen zustandekommen. Die Ergebnisse von SYZ (1926/27 b) und MANNING & MELCHIORI (1974) weisen darauf hin. Bei SYZ berichteten die Vpn bei Wörtern aus dem sexuellen Bereich keine Emotionen, obwohl diese PGR-Ausschläge hervorriefen;

dagegen gaben sie bei den Wörtern 'Vater', 'Mutter' an, Emotionen empfunden zu haben, obwohl kaum Ausschläge vorlagen. SYZ machte soziale Einstellungen, Normen und Vorurteile dafür verantwortlich.

Auch bei MANNING & MELCHIORI (1974) zeigte sich, daß sexuelle Items, trotzdem sie physiologisch stark aktivierend waren, als weniger aufregend (upsetting) beurteilt wurden als Gewaltitems, die physiologisch schwächere Reaktionen zur Folge hatten. Verblüffend ist die Übereinstimmung der Daten von SYZ (1926/27 b) und MANNING & MELCHIORI (1974), obwohl fast 50 Jahre zwischen den beiden Untersuchungen liegen.

Weitere Faktoren, von denen die Urteile abhängen, sind der Kontext und die Situation. Daß die unmittelbare Umgebung eines Wortes, also die vorausgehenden Wörter in einer Liste das Urteil mitprägen, hat AMSTER (1964) gezeigt. Je nach der Zusammensetzung der Wortliste (Bezugssystem) kommt es zu unterschiedlichen Ergebnissen. In einer heterogenen Liste fallen die Urteile extremer aus (Kontrasteffekt) als in einer homogenen Liste, die nur aus angenehmen oder nur aus unangenehmen Wörtern besteht. Setzt sich die Liste nur aus schlechten (bad) Wörtern zusammen, werden diese als nicht so unangenehm eingeschätzt als wenn sie sich in einer Liste befinden, in der auch gute (good) Wörter enthalten sind. Entsprechendes gilt für die guten Wörter. Diese werden in einer gemischten Liste positiver beurteilt als in der Nachbarschaft anderer angenehmer Wörter. Entscheidend ist, aus welchen Wörtern die Liste zusammengesetzt ist, ob aus extremen Stimuli oder aus sehr ähnlichen.

Auch die Situation, in der die Vp die Wörter beurteilen soll, kann eine Rolle spielen. Unter der Bedingung absichtliches Lernen wurden mittel schlechte Wörter schlechter beurteilt als beim inzidentellen Lernen. (AMSTER 1964)

Neben der Situation und der momentanen Verfassung muß vermutlich auch das Alter und das Geschlecht der Vp be-

rücksichtigt werden. Abhängig von den spezifischen Umwelt-
erfahrungen können dieselben Wörter für verschiedene Menschen
jeweils etwas anderes bedeuten.

Sicher spiegelt die subjektive Einschätzung nicht vollkommen
die Emotionalität von Wörtern wider. TRAXEL & BECHER (1957)
sind der Ansicht, daß die Korrelation zwischen Urteil und
Affekt nicht den Wert 1 erreichen kann.

Die bis jetzt aufgeführten Methoden kann man als 'subjektive'
Indikatoren der Emotion bezeichnen. Sie basieren auf einem
Urteil über den vermuteten oder empfundenen Affekt. Neben
diesen 'subjektiven' Indikatoren gibt es auch sogenannte
'objektive'. Hierzu zählen vor allem physiologische Reaktio-
nen auf Wörter, aus denen man auf einen emotionalen Faktor
zu schließen versucht.

2.2.2. Physiologische Reaktionen

Der am meisten verwendete Indikator ist die elektrodermale
Reaktion (EDR, PGR). Die Messung der EDR wurde schon zu
Beginn dieses Jahrhunderts zur Feststellung der Emotionali-
tät von Wörtern herangezogen, so z.B. von PETERSON & JUNG
(1907), WELLS & FORBES (1911), W. SMITH (1922/1970), WECHS-
LER (1925), SYZ (1926/27), JONES & WECHSLER (1928).

Als weitere physiologische Variablen wurden noch Atmung,
EEG und EKG verwendet.

Die anderen Maße wie Blutdruck und endokrine Veränderungen
eignen sich nicht für die Untersuchung kurzfristig wirksamer
Stimuli oder für die Untersuchung von Stimulusfolgen mit
kurzen Intervallen, da sie relativ 'träge' sind. (JANKE in
SCHÖNPFLUG 1969) Die Wirkung tritt entweder verspätet auf,
oder aber die Maße sind nicht sensibel genug, um schon mini-
male Änderungen in der Aktiviertheit zu erfassen, wie sie
durch die Darbietung von emotional besetzten Wörtern erzielt
werden.

Aus den physiologischen Reaktionen lassen sich gegenwärtig nur Aussagen über die Emotionalität vs. Nicht-Emotionalität eines Wortes oder über die Intensität der Erregung ziehen. Welche Emotion durch ein Wort ausgelöst wird, kann beim heutigen Stand der Forschung nicht durch physiologische Messung festgestellt werden.*

Atmung

Gemessen werden Veränderungen in der Atmung, z.B. das Anhalten der Atmung oder tiefes Einatmen. LEHMANN (1892, zit. nach PETERSON & JUNG 1907) stellte fest, daß angenehme Reize eine tiefere Atmung auslösen und unangenehmes Stimulusmaterial mehrere tiefe Atembewegungen hervorruft. Nach CRAMER (1968) zeigten manche Untersuchungen, daß die Atmung von der Emotionalität des Stimuluswortes betroffen wird (BOYD & VALENTINE 1953).

Die Atmung fand aber letztendlich keinen direkten Eingang in die Bestimmung der Emotionalität von Wörtern. Mehrere Gründe sprechen dagegen.

- Die Atmung ist bewußt beeinflufßbar.
- Aufgeregte Menschen atmen in einer Situation wie dem Experiment, das bei ihnen eine nervöse Spannung hervorruft, insgesamt unruhig.
- Die Atmung ist außerdem nicht sensibel genug, so daß die meisten Wörter hier keine Veränderung hervorrufen.

Als sinnvoll erscheint es aber, die Atmung zur Kontrolle mit aufzuzeichnen, da Husten, tiefes Einatmen, Gähnen und Lachen psychogalvanische Reaktionen auslösen.

* Zu überprüfen wäre, inwieweit die Pupillometrie zur Untersuchung von Wörtern herangezogen werden kann, da laut E. HESS (zit. nach 'Die Zeit' Nr.28, 4.7.1980) aus der Pupillometrie Aussagen über die Richtung der Emotion (angenehm oder unangenehm) möglich sind. Bei angenehmem Empfinden erweitern sich die Pupillen, bei Unlust verengen sie sich. Eventuell verspricht auch die Analyse der Stimmfrequenz beim Aussprechen von Wörtern für die Zukunft genauere Aufschlüsse über die damit verbundenen Emotionen.

EEG

EMRICH & HEINEMANN (1966) fanden Unterschiede in der Ausprägung des Alphasrhythmus bei neutralen und gefühlsbetonten Wörtern. Bei E-Wörtern zeigte sich eine vermehrte Alphaausprägung. WARREN & HARRIS (1975) differenzierten die Wörter danach, ob ein Alphablock auftrat oder nicht.

EKG

In der Untersuchung von EMRICH & HEINEMANN (1966) fiel die Pulsfrequenz bei gefühlsbehafteten Wörtern im unter-schweligen Bereich niedriger aus als bei neutralen Wörtern. Im Schwellenbereich zeigten sich keine Unterschiede.

Körperbewegung

Bei RAPAPORT (1977, S. 76) wird auch ein Versuch erwähnt, bei dem die gesamte Körperbewegung als Indikator verwendet wurde, um die Gefühlsbetonung von Wörtern zu messen (CASON 1934).

EDR

Der empfindlichste Indikator bleibt aber nach wie vor die elektrodermale Reaktion. Da in meiner Untersuchung dieses Maß zur Messung herangezogen wird, soll im folgenden Abschnitt ausführlicher auf diese Methode eingegangen werden.

Exkurs: Die elektrodermale Aktivität

Die Messung der elektrodermalen Aktivität ist eine der gebräuchlichsten und beliebtesten Methoden zur Erfassung der Intensität emotionaler Vorgänge oder der Aktivierung. Ein Grund dafür liegt in der hohen Sensibilität (leichten Ansprechbarkeit), in der relativ einfachen technischen Realisierung der Messung und in dem im Vergleich zu anderen psychophysiologischen Reaktionen niedrigen 'biologischen Grundrauschen' (HEYDEN 1979, S.83). Die elektrodermale Aktivität läuft völlig unbewußt ab, sie wird vom autonomen Nervensystem reguliert. Sie läßt sich mit großer Genauig-

keit messen und scheint nach W.SMITH (1922/1970) jede Art von emotionaler Erregung zu begleiten. W.SMITH schreibt weiter (S.25): "this phenomenon is by far the most delicate, reliable and quantitatively accurate method at present known for detecting and measuring emotional changes." In der Literatur finden sich viele enthusiastische Lobeshymnen. WECHSLER (1925, S.116/117) stellt fest: "the psychogalvanic reflex is primarily, if not exclusively, an index of affective tone." TRAXEL (1963, S. 249/250) zum selben Thema: "Die Eignung dieser Erscheinung als Indikator der Intensität emotionaler Vorgänge ist durch ein reichhaltiges empirisches Material gesichert."

Das Phänomen der psychogalvanischen Reaktion wurde zum erstenmal 1888 von FÉRÉ beschrieben, der einen schwachen Strom durch zwei Hautpunkte seiner Vpn fließen ließ und eine Abnahme des Hautwiderstandes bei emotionaler Stimulierung feststellte. 1890 erhielt TARCHANOFF ähnliche Resultate ohne Verwendung einer externen Stromquelle. Zwei Hautpunkte wurden durch ein Galvanometer verbunden und zeigten eine Potentialdifferenz. Reize verursachten zusätzliche Potentialänderungen. Die Bezeichnung 'psychogalvanischer Reflex' oder 'psychogalvanische Reaktion' (PGR) wurde 1909 von VERAGUTH eingeführt. GILDEMEISTER & ELLINGHAUS (1923, zit. nach HEYDEN 1979, S.83) prägten den neutraleren Begriff 'Hautgalvanische Reaktion' (HGR) oder 'galvanischer Hautreflex' (GHR). Im englischen Sprachraum ist die Bezeichnung GSR (galvanic skin response) oder EDR (electrodermal response)* gebräuchlich. Diese Bezeichnungen (PGR, HGR, GSR, EDR) betreffen alle die kurzfristige Änderung des Hautwiderstands, die nach Stimulierung auftritt. Neben der phasischen Reaktion kann man auch den Grundwert (z.B. den Basalwiderstand der Haut) oder das Grundniveau messen, das den Grad der allgemeinen Aktivierung anzeigt und sich je nach Aktivierung des Individuums verändern kann.

* deutsch: 'elektrodermale Reaktion'. Diese Bezeichnung wird im folgenden Text verwendet.

Bevor nun auf die physiologischen Grundlagen, die biologische Bedeutung, die Messung und die Auswertung der elektrodermalen Aktivität eingegangen wird, erfolgt an dieser Stelle ein Überblick über die heute übliche Terminologie. Diese Zusammenfassung basiert auf einem terminologischen Klärungsversuch von WALSCHBURGER (1975, S.515/516 und 1976, S.235/236, z.T. im Anschluß an LYKKEN & VENABLES 1971, S.657) und den Tabellen bei GROLL-KNAPP (in Schönpflug 1969, S.254) und HEYDEN (1979, S.84).

Terminologie

Als allgemeiner Ausdruck für hautelektrische Phänomene fungiert der Oberbegriff 'Elektrodermale Aktivität' (EDA). Je nach Art der Messung (ob ohne äußere Stromquelle oder mit, und hier ob mit Strom oder mit Spannung) und nach Art des Maßes (tonisch oder phasisch) existieren folgende Bezeichnungen:

EXOSOMATISCHE MESSUNG (Messung mit externer Stromquelle)

tonische Maße

Hautwiderstandsniveau oder Grundwiderstand der Haut (bei Messung mit Strom)	-	skin resistance level	SRL
Hautleitwertniveau oder Leitfähigkeit der Haut (bei Messung mit Spannung)	-	skin conductance level	SCL
Hautimpedanzniveau (bei Messung mit Wechselstrom niedriger Frequenz)	-	skin impedance level	SZL
Haut-Scheinleitwertsniveau (bei Messung mit Hilfe einer Wechselspannung)	-	skin admittance level	SAL

phasische Maße allgemeine Bezeichnungen: PGR, HGR, GHR, EDR, GSR

Hautwiderstandsreaktion Widerstandsänderung der Haut (kurzfristige Abnahme des Hautwiderstandes)	-	skin resistance response	SRR
Hautleitwertsreaktion oder Leitfähigkeitsänderung der Haut (kurzfristiges Anwachsen des Leitwerts)	-	skin conductance response	SCR

Hautimpedanzreaktion	-	skin impedance response	SZR
Haut-Scheinleitwertsreaktion	-	skin admittance response	SAR

ENDOSOMATISCHE MESSUNG (Messung ohne externe Stromquelle)

tonisches_Maß

Hautpotentialniveau	-	skin potential level	SPL
---------------------	---	----------------------	-----

phasisches_Maß

Hautpotentialreaktion oder Hautpotentialänderung	-	skin potential response	SPR
---	---	-------------------------	-----

Niveauewerte

Das Grundniveau (der Grundwiderstand SRL oder die Grundleitfähigkeit SCL) gilt als Maß für den Grad der allgemeinen Aktivierung. Die Hautleitfähigkeit ist hoch (der Hautwiderstand niedrig), wenn der Organismus wach, aufmerksam, angespannt oder stark aktiviert ist und niedrig (der Hautwiderstand entsprechend hoch), wenn der Organismus entspannt und ruhig ist. Es zeigt sich ein typischer Tagesverlauf ähnlich wie bei der Körpertemperatur. In der Nacht während des Schlafs ist die Hautleitfähigkeit niedrig, am Morgen beginnt sie zu steigen und erreicht mittags ihr Maximum, dann fällt sie wieder ab, zuerst langsam, dann schneller. (Für den Hautwiderstand gelten die reziproken Werte.)

Es zeigen sich große interindividuelle Unterschiede in den Grundwerten. Der Grund dafür ist das Hautsystem der jeweiligen Vp und ihre Aktivierung.

Reaktionswerte

Die phasischen Änderungen können durch äußere Reize, durch innere Stimulation ("schwer kontrollierbare zentrale Aktivität, z.B. plötzliches Einschließen subjektiv bedeutsamer Gedanken oder Vorstellungen" WALSCHBURGER 1975, S.516) oder durch Körperbewegungen, Husten und tiefes Einatmen hervorgerufen werden. Die Leitfähigkeit der Haut nimmt dabei kurzfristig zu (entsprechend sinkt der Hautwiderstand für kurze Zeit). Dies geschieht selbst bei Erregungen, die so

gering sind, daß sie die Vp kaum oder gar nicht wahrnimmt. Je intensiver die Reize, desto stärker die elektrodermale Reaktion. Die Stärke der EDR hängt ab von der subjektiven Bedeutung des Reizes (BRINGHAM 1943, zit. nach TRAXEL 1963). Die elektrodermale Reaktion ist sehr sensitiv, schon geringste Anlässe lösen sie aus, sie erfolgt relativ schnell (die Latenzzeit beträgt durchschnittlich eine bis drei Sekunden) und zeigt eine momentane kurze Aktivierung an, die für die Vp unbemerkt abläuft. WOODWORTH & SCHLOSBERG (1965) geben für visuelle Reize eine Reaktionszeit von 2.1 sec und für akustische eine von 1.7 sec an. Bei wiederholter Gabe eines Reizes kann eine Habituation erfolgen, der Reiz ruft dann keine EDR mehr hervor.* Eine Habituation tritt nur ein, wenn der Reiz keine Bedeutung mehr hat. Die Habituation besteht nicht für längere Zeit, nach einer Pause kann der Reiz wieder eine EDR auslösen.

Die EDR wird als Indikator von Aktivierung, Orientierungsreaktion und Emotionen angesehen. Zur Verdeutlichung seien eine Reihe von Reizen aufgezählt, die in der Untersuchung der EDR verwendet wurden. Die Skala reicht von lauten Geräuschen, Lichtblitzen, E-Schocks, Nadelstichen, Androhung von E-Schocks oder Nadelstichen (dies rief manchmal sogar höhere EDR als der tatsächliche Reiz hervor), Berührung der Vp (z.B. an die Nase greifen) bis zu Gerüchen, Bildern, Filmen, Wörtern, Bearbeiten von Aufgaben (z.B. Rechenaufgaben), Vorstellungen. Plötzliche und neuartige Reize rufen eine Orientierungsreaktion hervor.

* Bei einem Vorversuch der Verfasserin zeigte sich bei der Wiederholung eines Wortes keine Abnahme in der Reaktion, sondern sogar eine Zunahme. Das Wort 'Sohn' löste bei der Vp mit der Zeit immer stärkere emotionale Reaktionen aus. Ein Problem mit dem Sohn der Schwester wurde ihr immer deutlicher bewußt und dadurch belastender.

Physiologische Grundlagen

Über die physiologischen Grundlagen der elektrodermalen Phänomene existierten früher die unterschiedlichsten Auffassungen. Man machte z.B. die Blutzirkulation (STICKER 1897, zit. nach PETERSON & JUNG 1907), die Vasomotorik (FÉRE 1888, zit. nach GROLL-KNAPP in SCHÖNPFLUG 1969), die Muskelaktivität (SIDIS & NELSON 1910*) oder die Schweißsekretion (DARROW 1934*) verantwortlich.

D'ARSONVAL*, ein Physiker, der mit FÉRE zusammenarbeitete, betonte als erster die Bedeutung der Schweißsekretion. Heute weiß man, daß die Aktivität der Schweißdrüsen eine zentrale Rolle spielt, daß aber das Auftreten von Schweiß an der Körperoberfläche nicht notwendig ist. Bei Menschen, die von Geburt an keine Schweißdrüsen besitzen (HAIDER in SCHÖNPFLUG 1969)**, läßt sich keine elektrodermale Aktivität messen. Am besten meßbar ist das Phänomen der EDA an Stellen, die reich an Schweißdrüsen sind. Zwischen der Anzahl der aktiven Schweißdrüsen und den Hautwiderstandswerten bestehen nach den bisherigen Untersuchungen Korrelationen von .4 bis .9 (THOMAS & KORR 1957, zit. nach EDELBERG 1972 und LADER in LEVI 1975).

Die exakte Natur der Vorgänge ist noch nicht restlos geklärt. Man nimmt an, daß Änderungen der Permeabilität der für Ionen unterschiedlich durchlässigen Membranen verantwortlich sind, für die tonischen Änderungen die Membranen im Stratum germinativum, Stratum granulosum (HEYDEN 1979) und Stratum lucidum (HAIDER in SCHÖNPFLUG 1969, S.130), für die phasischen Änderungen die Membranen der Schweißdrüsen. Erstere ändern ihre Permeabilität nur langsam. Die präsekretorische Aktivität führt zu einer erhöhten

* zit. nach WOODWORTH & SCHLOSBERG 1965, S.141 f.

** Unter den etwa 130 Vpn meiner Untersuchung befand sich eine männliche Vp, die in der Hand keine Schweißdrüsen besaß und mit der deshalb der Versuch nicht durchgeführt werden konnte.

Permeabilität der Schweißdrüsenzellmembranen und dadurch zum Absinken des Hautwiderstands. Die Aktivität der Schweißdrüsen wird ausschließlich vom sympathischen Teil des autonomen Nervensystems innerviert. Acetylcholin ist dabei die Überträgersubstanz.

Biologische Bedeutung

Die elektrodermale Reaktion wird als vom vegetativen Nervensystem ausgelöste Bereitstellungsreaktion angesehen. Sie dient der vegetativen Anpassung an 'Notfallsituationen' (WALSCHBURGER 1976, S.241). Die Fähigkeit, auf bestimmte Veränderungen in der Umwelt sinnvoll zu reagieren, soll dadurch unterstützt werden.

REIZZUWENDUNGSREAKTION

Durch die Schweißdrüsenaktivität wird die Sensitivität der Haut erhöht. KATZ (1925, zit. nach EDELBERG 1972) zeigte, daß feuchte Haut sensibler ist als trockene Haut. Durch die Feuchtigkeit werden nach DARROW (1936, zit. nach EDELBERG 1972) Greifen und taktile Manipulation erleichtert. Für die taktile Anpassung ist ein mittlerer Grad an Feuchtigkeit optimal.

ABWEHRREAKTION

Die schwitzende Haut ist weniger anfällig gegen Risse und Abschürfungen. Durch das Schwitzen wird der Organismus vor Verletzungen geschützt. Nach EDELBERG (1972) ist es äußerst schwierig, in eine schlüpfrige Haut zu schneiden. Bedrohliche Situationen rufen diese Schutzreaktion hervor.

THERMOREGULATION

Eine weitere Bedeutung der Schweißdrüsenaktivität liegt in der thermoregulatorischen Anpassung. Die Kühlung des Körpers wird vorbereitet, wenn eine Zunahme von Stoffwechsel- und motorischer Aktivität erwartet wird, was z.B. bei 'Kampf' oder 'Flucht' der Fall ist. Wenn sich unser Körper auf eine mögliche Gefahr einstellt, so verengen sich die Blutgefäße der Haut. Der Blutdruck steigt und der Blut-

fluß verlangsamt sich, wodurch bei eventuellen oberflächlichen Verletzungen der Blutverlust möglichst gering gehalten wird. Weil sich dadurch aber die Körpertemperatur erhöht, muß sie durch Verdunstungskälte wieder gesenkt werden. Diese wichtige, lebenserhaltende Funktion übernehmen die Schweißdrüsen. Ob alle Schweißdrüsen in gleicher Weise für die Thermoregulation zuständig sind, ist noch nicht endgültig geklärt. Während KUNO (1956, zit. nach EDELBERG 1972) behauptet, daß dies für die Schweißdrüsen von Handflächen und Fußsohlen nicht zutrifft, stellt WILCOTT (1963, zit. nach VENABLES & MARTIN 1967) im Gegensatz dazu fest, daß alle Schweißdrüsen, nicht nur die des Rumpfes an der Thermoregulation beteiligt sind. Nicht nur die Schweißdrüsen an den Handinnenflächen und Fußsohlen, sondern auch die anderen Schweißdrüsen des Körpers können durch emotionale Reize aktiviert werden.

Eine weitere mögliche Bedeutung der Schweißsekretion wird darin gesehen, daß Geruchsreize ausgesendet werden, die das Auffinden und Erkennen von Artgenossen erleichtern sollen.

Messung

Die Messung der elektrodermalen Aktivität wird folgendermaßen vorgenommen. Abgeleitet wird von Hautstellen, an denen Schweißdrüsen gehäuft auftreten, wie die Handinnenfläche, die Fingerinnenseite oder die Fußsohlen, die aber seltener verwendet werden. Nach WALSCHBURGER (1975, S.516) sprechen vor allem diese Schweißdrüsen auf sensorische Reize und emotionale Veränderungen an im Gegensatz zu "den primär thermoregulatorisch tätigen Schweißdrüsen des Rumpfes". Für die Ableitung von den Fingern benötigt man sehr kleine Elektroden, die meist an den mittleren Gliedern von Mittel- und Ringfinger befestigt werden. (siehe Abb. 1a) Für die Ableitung von der Handinnenfläche spricht der bessere Sitz der Elektroden, die hier etwas größer sein können, was eine zuverlässigere Messung erlaubt. Die Elektroden-

applikation erfolgt hier so, daß eine der beiden Elektroden auf dem Daumenballen (thenar) und die andere auf dem gegenüberliegenden proximalen Drittel des Kleinfingerballens (hypothenar) angebracht wird. (siehe Abb. 1b)

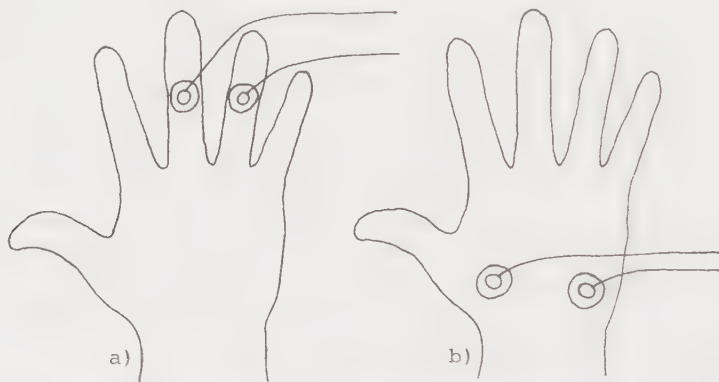


Abb. 1: Ableitungsstellen zur Messung der EDA

a) von der Fingerinnenseite b) von der Handinnenfläche

Wichtig ist ein guter, gleichmäßiger Kontakt der Elektroden mit der Haut, um Artefakte zu vermeiden. Die früheren Plat-tenelektroden waren in dieser Hinsicht äußerst problematisch, da hier ein gleichbleibender Kontakt kaum gewährleistet war. Durch minimale Bewegungen oder Zittern der Hand veränderte sich die Kontaktfläche und rief dadurch Veränderungen in der abgeleiteten Kurve hervor. WALSCHBURGER (1975) empfiehlt Ag/AgCl-Elektroden, da sie geringe Polarisationsneigung zeigen.

Die Elektroden werden an das Meßgerät angeschlossen. Bei der endosomatischen Methode, also der Messung ohne äußere Stromquelle, werden die Elektroden mit einem Galvanometer verbunden. Diese Art der Messung (die Messung des Hautpotentials) wird als weniger zuverlässig erachtet und deshalb auch nicht so häufig verwendet wie die exosomatische Methode. Im folgenden wird aus diesem Grund nur die exosomatische Methode behandelt. Eine ausführliche Darstellung der Messung des

Hautpotentials findet sich bei VENABLES & MARTIN (1967).

Bei der exosomatischen Methode unterscheidet man zwei Arten von Messung, entweder die Messung des Hautwiderstands, die vor allem früher durchgeführt wurde, oder die Messung der Hautleitfähigkeit. Bei ersterer wird ein konstanter Strom durch die Haut geschickt; die Stromdichte soll dabei $10\mu\text{A}/\text{cm}^2$ nicht übersteigen. Die klassische Messung des Hautwiderstandes wurde mit der Wheatstonschen Brücke durchgeführt, die aber Nachteile in sich birgt. Gegenwärtig verwendet man meistens Geräte mit Kompensationsschaltungen, bei denen der Signalverlauf in absoluten Einheiten eichbar ist (LANC 1977). Zur Messung der Hautleitfähigkeit wird eine konstante Spannung angelegt, die nicht mehr als 0,5 Volt betragen soll.

Heute zeigt sich eine Tendenz zur Bevorzugung der Messung der Hautleitfähigkeit (LADER in LEVI 1975; WALSCHBURGER 1975; HEYDEN 1979; SCHANDRY, persönliche Mitteilung 1979). Physiologische Gründe sprechen dafür, weil bei der Hautleitfähigkeitsmessung "die Hautareale nur entsprechend ihrer Leitfähigkeit Strom führen, während bei der Widerstandsmessung bestimmte Gewebeteile ... unter Umständen überlastet werden können" (HEYDEN 1979, S.87). Ebenso sind bei der U_K -Methode (Messung mit konstanter Spannung) Verstärkungs- bzw. Empfindlichkeitsverstellungen seltener notwendig (WALSCHBURGER 1975).

Der Wert, den man bei der Messung erhält, ist abhängig vom Typ und von der Größe der Elektroden (je kleiner die Elektrode ist, desto größer ist der Widerstand), von der Beschaffenheit der Hautregion (Anzahl der Schweißdrüsen), der Raumtemperatur und dem momentanen Aktivierungsgrad der V_p .

Der Widerstand wird in Ohm gemessen, die Leitfähigkeit in micromho oder Siemens (S). Widerstand und Hautleitfähigkeit sind umgekehrt proportional zueinander. 1 micromho entspricht 1 000 000 Ohm.

$$1 \text{ Siemens} = 1 \text{ mho} = \frac{1}{\text{Ohm}}$$

BOUCSEIN & HOFFMANN (1979, zit. nach HEYDEN 1979) fanden in ihrer Untersuchung keine Unterschiede zwischen SRR und SCR in bezug auf die Amplitude und die Fläche, aber bzgl. der Recoveryzeit, die bei der Leitfähigkeitsmessung signifikant kürzer war.

Aufgezeichnet wird der Verlauf der elektrodermalen Aktivität mittels eines Schreibers. Früher geschah das z.B. mit einem Lichtstrahl, der auf lichtempfindliches Papier geworfen wurde. Die Aufzeichnung mußte daher in einem dunklen Raum erfolgen, oder ein VL fuhr die Lichtspuren mit einem Tintensiphon nach (wie z.B. bei WECHSLER & JONES 1928). Eine verbesserte Technik ermöglicht heute eine problemlose Registrierung z.B. mit thermischen Schreibern oder Druckkopierschreibern, bei denen die Signale sauber auf das Papier übertragen werden im Gegensatz zu den Tintenschreibern.

Aufgezeichnet werden am häufigsten die phasischen Änderungen der elektrodermalen Aktivität, seltener der Verlauf der tonischen Werte. Bei den modernen Polygraphen ist die gleichzeitige Registrierung beider Informationen möglich.

Auswertung

Zur Beschreibung der phasischen Veränderungen kann man verschiedene Reaktionsparameter heranziehen: die Amplitude einer EDR, ihren Anstiegswinkel, verschiedene Zeitmaße, z.B. die Latenzzeit, die An- und Abstiegszeit, die Halbwertzeit, die Recoveryzeit und die Gesamtdauer, verschiedene Flächenmaße und schließlich die Frequenz (Anzahl der EDR pro Zeiteinheit). Abbildung 2 zeigt den typischen Verlauf einer EDR und veranschaulicht die verwendbaren Maße.

Die Amplitude gibt die Höhe der Reaktion an. Sie ist die Differenz zwischen dem Wert zu Beginn der Reaktion und dem höchsten Punkt des Ausschlags (ΔR oder ΔI). Die Amplitude ist die am häufigsten verwendete Variable.

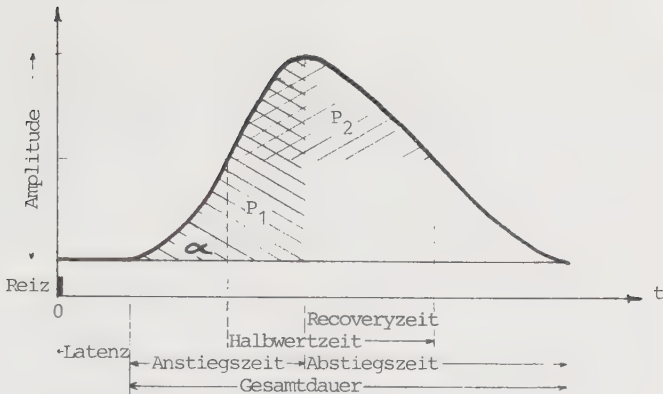


Abb. 2: Typischer Verlauf einer elektrodermalen Reaktion mit den verwendbaren Maßen.

Der Tangens des Anstiegswinkels α gibt die Steilheit des Anstiegs einer EDR an.

Durch die Latenzzeit wird die Zeit zwischen dem Reiz und dem Beginn der Reaktion ausgedrückt. Die Anstiegszeit geht von Beginn der Reaktion bis zum höchsten Punkt des Ausschlags, die Abstiegszeit vom höchsten Punkt bis zum Ende der Reaktion. Anstiegszeit und Abstiegszeit zusammen ergeben die Gesamtdauer. Die Halbwertzeit wird auf einem etwas umständlicheren Weg ermittelt. Sie gibt die Zeitdauer zwischen den Punkten auf der Anstiegs- und der Abstiegsseite an, bei denen die halbe Amplitudenhöhe vorliegt. Die Recoveryzeit wird vom Zeitpunkt des maximalen Ausschlags bis zu dem Punkt gerechnet, an dem die halbe Amplitudenhöhe erreicht ist.

Soll die Auswertung mit Hilfe von Flächenmaßen erfolgen, so kann man entweder die Fläche innerhalb der Anstiegszeit (P_1) verwenden oder die Fläche über der Halbwertzeitlinie (P_2).

Über die Beziehung der einzelnen Kennwerte untereinander gibt es noch kaum Aussagen. Amplituden- und Zeitmaße korrelieren nicht oder nur schwach miteinander.

Bei einer einzelnen EDR ist die Auswertung relativ einfach. Schwieriger wird es erst, wenn sich mehrere elektrodermale Reaktionen überlagern.* In diesem Fall läßt sich die Amplitude auf zwei Arten bestimmen. (siehe Abb. 3 a und b)

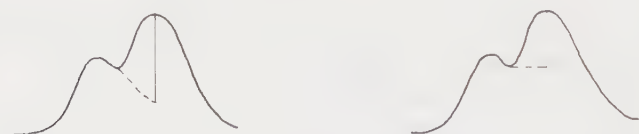


Abb. 3: Amplitudenbestimmung bei mehrphasigen Ausschlägen

- a) Man verlängert die erste absteigende Linie und fällt auf sie die Amplitude. b) Die Amplitude wird nur bis zu dem Punkt des Anstiegs gemessen.

Bei der Methode a) ergibt sich eine größere Amplitude als bei b). Laut EDELBURG (1967, zit. nach WALSBURGER 1976) ist die zweite Methode (b) vorzuziehen.

Die absoluten Meßwerte der EDR sind bei unterschiedlichem Grundniveau nicht direkt vergleichbar (Ausgangswertgesetz!). Die Größe der Amplitude hängt u.a. vom Grunderregungsniveau ab. Bei einer stark aktivierten Vp wird ein neuer Reiz ihre physiologische Erregung nur noch wenig steigern können (Deckeneffekt). Sie wird niedrigere Ausschläge aufweisen als eine Vp, die weniger stark erregt ist und ihre obere Grenze noch nicht erreicht hat.

Die Beziehung zwischen Grundniveau und phasischer Reaktion ist noch nicht geklärt. Von manchen Autoren (zit. nach GROLL-KNAPP in SCHÖNPFLUG 1969) wird eine lineare Beziehung vertreten wie z.B. von SEWARD et al. (1935), ELLSON et al. (1952) und DAVIS et al. (1955). FREEMAN & KATZOFF

* Die mehrphasigen Ausschläge (doppelte oder mehrere sich überlagernde) kommen laut VERAGUTH nach starken Reizen vor, laut WECHSLER (1925) zeigen sie sich öfters bei gefühlsbetonten Menschen, und PETERSON & JUNG (1907) nehmen sie als Indikator für große emotionale Labilität.

(1942) fanden eine U-förmige Beziehung mit den stärksten Reaktionen bei einem mittleren Grundniveau. Laut WALSCH-BURGER (1976, S.257) ist die tonisch-phasische Abhängigkeit bei der Verwendung von Leitwerteinheiten als geringer anzusehen ("unter Annahme des Parallel-Leitwert-Modells der Haut") als bei Widerstandseinheiten.

Da die absoluten Meßwerte bei unterschiedlichem Grundniveau nicht miteinander verglichen werden können, versucht man, die absoluten Werte in relative umzuwandeln. Es gibt verschiedene Transformationen. Meistens werden prozentuale Umwandlungen vorgenommen. Dabei wird die Amplitude der Hautwiderstandsreaktion (ΔR) durch den Grundwiderstand (R) geteilt ($\Delta R/R$). Das gleiche Verfahren gilt für die Leitfähigkeitsreaktion ($\Delta I/I$). Eine andere Methode ist die Umwandlung der Werte in Logarithmen, wobei die Werte der Amplitude von SCR oder SRR logarithmiert werden. HAGGARD (1949) spricht sich für eine logarithmische Transformation der Hautleitfähigkeitsreaktion aus (log conductance change). Die Verarbeitung der Werte bei W.SMITH (1922/1970) sieht folgendermaßen aus: die Amplitude der HGR wird durch den Mittelwert aller Amplituden der Vp dividiert:

$$\frac{\Delta R}{\bar{X}_{\Delta R}}$$

TRAXEL (1960) findet, daß die Quadratwurzel der relativen Widerstandsänderung ($\sqrt{\Delta R/R}$) der Gefühlsstärke proportional ist. Ungeklärt ist bisher, welche Transformation die adäquateste ist.

EDELBERG (1972) hält die Korrektur der absoluten Werte nicht immer für sinnvoll. Solange man innerhalb einer Vp bleibt, die kein stark verändertes Grundniveau während des Versuchs aufweist, kann man die absoluten Werte verwenden. Bei einer Zusammenfassung der Werte mehrerer Individuen oder einem Vergleich ist eine Umwandlung nötig. Hierfür konnte aber noch keine verbindliche Vorgehensweise formuliert werden.

Es gibt große individuelle Unterschiede in der Reagibilität des elektrodermalen Systems. Dies hängt u.a. von der Beschaffenheit der Haut ab, von der Morphologie und Funktion der Schweißdrüsen, deren Anzahl von Mensch zu Mensch verschieden ist. Zu beachten ist auch die unterschiedliche Ansprechbarkeit der physiologischen Systeme innerhalb einer Vp (LADER in LEVI 1975, FAHRENBURG in SCHÖNPFUG 1969): danach reagieren manche Menschen z.B. stärker mit einer Veränderung des Blutdrucks oder der Pulsfrequenz als mit einer Aktivierung des SchweißdrüSENSYSTEMS.

Elektrodermale Aktivität findet sich auch bei Kindern und Neugeborenen (JONES 1930). JONES, der Kleinkinder im Alter von drei Monaten an untersuchte, stellte kaum Unterschiede in der elektrodermalen Aktivität im Vergleich mit Erwachsenen fest.

Um möglichst exakte Meßergebnisse zu erzielen, muß eine Reihe von Versuchsbedingungen genau beachtet werden:

- Es dürfen keine störenden Reize auftreten.
- Der Versuch muß in einem ruhigen Raum stattfinden.
- Die Vp sollte möglichst ruhig und entspannt sitzen.
- Die Raumtemperatur sollte konstant und nicht zu niedrig sein. (20° C)
- Die Vp darf keine zu kalten Hände haben.*

Bei der ipsativen Betrachtungsweise ergeben sich keine Probleme bei der Interpretation der Werte. Schwieriger ist ein Vergleich über Individuen. Daß eine Vp höhere Amplitudenwerte hat, muß nicht bedeuten, daß für sie die Reize stärker aktivierend waren.

* Bei kalten Händen lassen sich wegen der kaum vorhandenen Aktivierung der Schweißdrüsen keine Reaktionen oder nur schlechte Ergebnisse beobachten. Ein solches Mißgeschick traf W.SMITH (1922/1970), der einen Teil seiner Versuche bei kaltem Wetter und gleichzeitigem Kohlenstreik durchführte.

2.2.2.1. Messung der emotionalen Intensität von Wörtern mit Hilfe der EDR

Im einfachsten Fall werden die Wörter nur in zwei Kategorien aufgeteilt, a) in emotionale, das sind solche, die eine elektrodermale Reaktion hervorrufen, b) in neutrale, diese Wörter lösen keine EDR aus.

Bei den Wörtern, die eine EDR bewirken, kann man eine weitere Differenzierung hinsichtlich der emotionalen Intensität anhand der Größe der EDR-Amplitude vornehmen. Dazu werden die Wörter nach der Höhe ihrer EDR-Amplitude in eine Rangreihe gebracht (W.SMITH 1922/1970, TRAXEL & BECHER 1957). W.SMITH gehörte mit zu den ersten, die die emotionale Tönung der Wörter anhand der EDR-Ausschläge beurteilten. Von 100 Wörtern, voraus 5 Übungswörter, die nicht berücksichtigt wurden, erhob er die EDR-Werte von 50 Vpn. Anhand der Größe der Amplituden (aller Vpn) erstellte er eine Rangreihe, wobei Rang 1 das Wort mit dem stärksten Ausschlag bekam und den Rang 100 das Wort mit der kleinsten Amplitude. (siehe Tabelle 1)

Wort	EDR	Wort	EDR	Wort	EDR	Wort	EDR
1. Kiss	72.8	26. Wine	30.9	51. Street	24.9	76. Try	20.0
2. Love	59.5	27. Luck	30.8	52. Beat	24.6	77. Plum	20.0
3. Marry	58.5	28. Green	30.4	53. Carry	24.5	78. Village	19.9
4. Divorce	50.8	29. Ask	30.0	54. Wait	24.4	79. Rich	19.9
5. Name	48.7	30. Make	29.9	55. Speak	24.3	80. Salt	19.8
6. Woman	40.3	31. Pity	29.7	56. Box	23.9	81. Bird	19.6
7. Wound	38.0	32. Choice	29.7	57. Nasty	23.6	82. Bread	19.6
8. Dance	37.4	33. Dress	28.5	58. Jump	23.5	83. Old	19.3
9. Afraid	36.8	34. Wicked	28.4	59. Paper	23.2	84. Cow	19.0
10. Proud	36.7	35. Dead	27.6	60. Lamp	23.1	85. Bring	19.0
11. Habit	36.6	36. Sing	27.6	61. Cold	23.0	86. Clean	18.8
12. Money	35.6	37. Horse	27.1	62. Long	22.7	87. Ink	18.7
13. Fight	35.0	38. Evil	27.0	63. Go	22.6	88. Sheet	18.6
14. Child	35.0	39. Doctor	26.9	64. Cook	22.3	89. Table	18.5
15. State	34.8	40. Stalk	26.2	65. Yellow	22.2	90. Work	18.3
16. Despise	34.7	41. Book	26.1	66. Chair	21.7	91. Carrot	18.2
17. War	34.1	42. Travel	25.9	67. Finger	21.5	92. Bury*	18.0
18. Family	33.6	43. Sick	25.8	68. Sad	21.4	93. Hunger	17.9
19. Happy	33.4	44. Bag	25.8	69. Tree	21.2	94. White	17.8
20. Pray	33.1	45. Water	25.6	70. Needle	21.1	95. Glass	17.6
21. Worry	33.0	46. Home	25.4	71. Blue	20.6	96. Give	16.7
22. Insult	32.5	47. Big	25.3	72. Ship	20.5	97. Flower	16.1
23. Friend	32.2	48. Bed	25.2	73. Motor	20.4	98. Pond	15.5
24. Head	31.7	49. Silly	25.2	74. Frog	20.2	99. Pencil	15.4
25. Angry	31.5	50. Brother	25.2	75. Walk	20.1	100. Swim	14.2

Tabelle 1: Rangreihe der Wörter nach Höhe der EDR-Amplitude
(W.SMITH 1922/1970, S.38).

* Dieses Wort wurde häufig als 'Berry' verstanden.

Meist erfolgt aber nur eine Aufteilung in zwei Intensitätsgrade (stärker vs. schwächer erregend). Die Gruppen können dabei auf verschiedene Arten gebildet werden. Die Wörter, deren Amplitude über dem Median liegt, fallen in die 'stärker emotionale' Wortgruppe, die mit einem Amplitudenwert unter dem Median werden als schwächer emotional angesehen (SCHÖNPFLUG & BEIKE 1964). Von ihren 24 Wörtern bezeichneten sie die 12 mit größerer Amplitude als Wörter 'hoher Ich-Bezogenheit', die 12 mit kleinerer Amplitude als Wörter 'niedriger Ich-Bezogenheit'. Möglich ist es auch, nur die Wörter am Anfang und am Ende der Rangreihe zu verwenden (KLEINSMITH & KAPLAN 1963, JONES & WECHSLER 1928). KLEINSMITH & KAPLAN (1963) verwendeten für ihre Gruppenbildung der 'high-arousal items' und 'low-arousal items' nicht alle dargebotenen acht Wörter*, sondern nur die ersten drei und die letzten drei Wörter der Rangreihe (der jeweiligen Vp). Seltener erfolgt eine Aufteilung in drei oder mehr Intensitätsstufen.

Die Bestimmung der Emotionalität der Wörter kann entweder ipsativ (KLEINSMITH & KAPLAN 1963, WECHSLER 1925, W.SMITH 1922/1970) oder normativ (W.SMITH 1922/1970, TRAXEL & BECHER 1957, SCHÖNPFLUG & BEIKE 1964) geschehen. Im ersten Fall werden die individuellen Amplitudenwerte in eine Rangreihe gebracht, so daß für jede Vp eine eigene Rangreihe entsprechend ihren persönlichen Reaktionen entsteht, die nur für sie gilt. Bei der normativen Auswertung werden pro Wort die Amplitudenwerte aller Vpn zusammengefaßt, und nur eine einzige Rangreihe anhand der Gruppenwerte gebildet.**Diese Daten gelten dann für die Population, aus der die Stichprobe stammt.

* kiss, rape, vomit, exam, dance, money, love, swim

** Hier besteht das Problem, wie man die absoluten Amplitudenwerte transformiert. Bei unterschiedlichem Grundniveau sind die absoluten Werte nicht vergleichbar, weshalb es zu einer Verzerrung der Daten kommen könnte.

2.2.2.2. Reliabilität

W.SMITH (1922/1970) erfaßte bei sechs Vpn sechsmal (an verschiedenen Tagen) die elektrodermalen Reaktionen auf 100 Wörter, die jeweils in unterschiedlicher Reihenfolge dargeboten wurden. 80 der 100 Wörter stammten aus der früheren Liste (siehe S.43, Tabelle 1). Die 20 erregendsten Wörter wurden durch 20 neutralere ersetzt. Es ergaben sich folgende Korrelationen (S.113/114):

Vp	r_{12}	$r_{(123)(456)}$
1	.69	.98
2	.39	.72
3	.18	.70
4	.44	.60
5	.35	.68
6	-.01	.42

Tabelle 2: Korrelationen der ersten mit der zweiten Messung (r_{12}) und Korrelationen der Durchschnittswerte aus den ersten drei Messungen mit den Durchschnittswerten aus den folgenden drei Messungen ($r_{(123)(456)}$).

Bei den Korrelationen der Durchschnittswerte einer Vp mit den Durchschnittswerten einer anderen Vp kam es nur zu sehr niedrigen Korrelationskoeffizienten. W.SMITH schließt daraus, daß im allgemeinen die Korrelation der Individuen mit sich selbst ungefähr .60 bis .70 (S.116) beträgt und daß die Korrelation verschiedener Individuen nicht größer als .20 (S.117) ausfällt. Diese niedrigen Korrelationskoeffizienten gelten aber nur für diese Liste, in der die allgemein am stärksten erregenden Wörter fehlen. Es gibt Wörter, auf die die Mehrzahl der Vpn in ähnlicher Weise reagiert (die Korrelationen fielen hier größer aus), und es gibt Wörter, die individuell sehr unterschiedlich aufgefaßt werden.

TRAXEL & BECHER (1957) führten mit neun ihrer 19 Vpn nach drei Wochen eine Wiederholungsmessung durch. Die Korrela-

tion für die Stabilität der Messungen der Größendifferenz des Widerstands betrug .64 (bei 10 Wortpaaren).

Übereinstimmung bei verschiedenen Vpngruppen

JONES & WECHSLER (1928) verwendeten 16 Wörter aus der Liste von W.SMITH (1922), acht davon mit niedrigem Rangplatz (= emotionale Wörter) und acht andere mit hohem Rangplatz (= neutrale Wörter).^{*} Sie erhoben bei 29 Vpn die EDR-Ausschläge auf diese Wörter und erhielten eine ähnliche Reihenfolge wie W.SMITH. Die Rangkorrelation beträgt .90 (S.608).

SCHÖNPFLUG & BEIKE (1964) übernahmen die 20 Wörter aus der Untersuchung von TRAXEL & BECHER (1957), fügten vier weitere hinzu und führten EDR-Messungen durch. Der Vergleich ihrer Daten mit der Rangreihe der Werte bei TRAXEL & BECHER^{**} zeigt bis auf zwei Ausnahmen eine gute Übereinstimmung. Das Wort 'Wetter', das bei TRAXEL & BECHER an 7.Stelle steht, rief bei ihren Vpn eine im Vergleich zu den anderen Wörtern kleinere Amplitude hervor, so daß es hier zu den neutralen zählt. Das Wort 'Überfall', auf Rangplatz 14 bei TRAXEL & BECHER, fällt bei ihnen aber unter die stärker erregenden Wörter.

* emotionale Wörter: 1. kiss, 2. love, 3. marry, 6.woman,
8. dance, 9. afraid, 12. money, 22.insult
neutrale Wörter: 100. swim, 99. pencil, 97. flower, 95. glass
94. white, 91. carrot, 90. work, 89. table.

^{**} Rangreihe der Wörter bei TRAXEL & BECHER (S.298) (mit den Durchschnittsgrößen der Hautwiderstandsänderungen d/R):

1. Examen	0,75%	11. Gruß	0,25%
2. Heirat	0,70%	12. Gedanke	0,21%
3. Liebe	0,60%	13. Schnupfen	0,19%
4. Kuß	0,60%	14. Überfall	0,16%
5. Angst	0,45%	15. Blume	0,16%
6. Tanz	0,40%	16. Salz	0,11%
7. Wetter	0,36%	17. Buch	0,10%
8. Erhängen	0,33%	18. Tee	0,10%
9. Erblindung	0,30%	19. Gras	0,09%
10. Baum	0,26%	20. Beleidigung	0,09%

2.2.2.3. Probleme bei der EDR-Messung

Hautwiderstandsänderungen oder Hautleitwertsreaktionen können durch andere emotionale Vorgänge, die mit der Wirkung des Reizwortes in keiner unmittelbaren Verbindung stehen, mitbestimmt werden. Vor allem zu Beginn des Versuchs braucht die Vp Zeit, sich zu orientieren. Während der ersten Wortdarbietungen treten infolge einer Erwartungsspannung unverhältnismäßige starke EDR-Ausschläge auf, sogenannte 'Orientierungsreaktionen'. Nach durchschnittlich fünf bis sechs Wörtern hat sich die Vp an die Situation angepaßt. Bei den ersten Wörtern ist es nicht möglich, den emotionalen Gehalt zu bestimmen, da die Orientierungsreaktionen stören; sie dürfen deshalb nicht gewertet werden.

JONES & WECHSLER (1928) zogen aus ihrer Untersuchung den Schluß, daß wenigstens fünf Übungswörter vorausgeschickt werden müssen. In ihrer Untersuchung hatten sie nur ein Übungswort vorgegeben. Sie boten 16 Wörter (4 neutrale, 4 emotionale, 4 neutrale und 4 emotionale) in zwei Reihenfolgen dar: die eine Hälfte der Vpn erhielt die Wortreihe vorwärts, die andere rückwärts. Je nach der Reihenfolge ergaben sich folgende Werte:

	4 neutr.	4 emot.	4 neutr.	4 emot.
Reihenfolge $\longrightarrow$	6.19	6.88	3.09	8.06
Reihenfolge $\longleftarrow$	1.40	6.44	3.62	9.25

Tabelle 3: Mittlere Amplitudenwerte* der beiden Darbietungsrichtungen.

Wenn die vier neutralen Wörter am Anfang stehen, bewirken sie hohe Ausschläge, werden sie erst am Ende präsentiert, rufen sie nur noch geringe Reaktionen hervor. Anders die emotionalen Wörter: sie rufen sowohl am Anfang als auch

* Die Durchschnittswerte wurden aus Tabelle II (JONES & WECHSLER, S.609) ermittelt.

am Ende der Liste hohe Ausschläge hervor.

Der Fehler, keine Übungswörter zu verwenden, ist sehr häufig begangen worden. Auch ein oder zwei Übungsbeispiele sind noch zu wenig. Ebenso falsch ist die Verwendung von Übungsstimuli, die anderer Art sind wie die Testitems (also z.B. Bilder anstelle von Wörtern). Die daraus resultierenden Werte für die Anfangswörter spiegeln nicht das eigentliche emotionale Erregungspotential wider.

Zu den Einflußgrößen, auf die bei der Arbeit mit der EDR zu achten ist, zählt die Habituation. Je länger der Versuch dauert, desto kleiner kann die EDR ausfallen, im Extremfall sogar überhaupt nicht mehr auftauchen. Zu berücksichtigen ist deshalb, daß die Stellung eines Reizes in der Reihenfolge nicht gleichgültig ist.

Das Problem, inwieweit ein vorausgehendes Wort ein nachfolgendes beeinflußt, spielt sowohl beim subjektiven Urteil als auch bei der EDR-Messung eine Rolle. JONES & WECHSLER (1928) berichten von einer Übertragung. Wenn das Wort 'glass' nach dem Wort 'woman' erschien, rief es einen höheren Ausschlag hervor als wenn es vor 'woman' und nach 'flower' dargeboten wurde.

Wie schon erwähnt ist die EDR sehr sensibel. Daher ist jede Art von Störung in der Umgebung zu vermeiden, ebenso sollte sich die Vp so ruhig wie möglich verhalten. Ist dies nicht der Fall, entstehen EDR-Ausschläge, die die Ergebnisse verfälschen.

Schließlich sei noch auf das Problem der Reaktivität hingewiesen. Da die Vp an Apparate angeschlossen wird, ist ihr ständig bewußt, daß an ihr Messungen vorgenommen werden, was wiederum einen Einfluß auf ihre Reaktionen haben kann.*

* Dieser Störfaktor ließe sich durch physiologische Messungen verhindern, bei denen die Vp nicht direkt mit den Geräten in Kontakt kommt. Die Feststellung der Hautdurchblutung mit Infrarotkammeras scheint ein Schritt in diese Richtung zu sein.

2.2.2.4. Validität

Ist sichergestellt, daß die elektrodermalen Reaktionen alleine auf die Wörter zurückzuführen sind (erkennbar an der Latenzzeit und bei Ausschluß von Störungen), so kann man davon ausgehen, daß das spezifische Wort eine Aktivierung hervorgerufen hat. Wenn das Wort nicht am Anfang der Liste steht, gibt die Aktivierung das emotionale Erregungspotential wieder. In den Untersuchungen, in denen keine Übungswörter verwendet wurden, liegt eine Konfundierung vor.

KAPLAN & KAPLAN (1968, S.291) erachten Wörter, die zu einer hohen elektrodermalen Reaktion führen, nicht in jedem Fall für emotional, da nach ihrer Meinung keine unbedingte Äquivalenz zwischen Aktivierung und Emotion vorliegt. Sie sprechen zwar von der Möglichkeit einer starken Emotionalität ("they may be high in 'emotionality'"), meinen aber, daß das nicht immer so sein muß.*

Wird die EDR-Messung exakt und sauber durchgeführt, so läßt sich mit ihr zuverlässig feststellen, welches Wort bei der Vp eine Erregung hervorruft und welches neutral ist. Wodurch ein Ausschlag bewirkt wird, vor allem bei allgemein als neutral angesehenen Wörtern (die Mehrzahl der Vpn hat hier keine Ausschläge), ist an den EDR-Daten nicht ablesbar. Die Bedeutung der EDR ist nur durch das Urteil der Vp eruierbar (TRAXEL & BECHER 1957). Dazu einige Beispiele:**

* In den meisten Studien von KAPLAN et al. wurde auf Übungswörter verzichtet, deswegen spiegeln ihre EDR-Werte nicht rein den emotionalen Aspekt der Wörter wider, sondern eine Aktivierung, die mit durch die Position bewirkt ist.

** Bei einem Versuch zum verbalen Konditionieren (Verfasserin et al.) wurden 'neutrale' Wörter mehrmals in zufälliger Reihenfolge dargeboten. Dabei fielen bei einigen Vpn bestimmte Wörter auf, die fast jedesmal einen Ausschlag produzierten. Bei der anschließenden Befragung konnten die Gründe für die Aktivierung ermittelt werden.

- Den starken Ausschlag beim Wort 'Jacke' führte eine Vp darauf zurück, daß sie dabei verärgert an ihre neue Lederjacke dachte, die durch starken Regen Flecken bekommen hatte.
- Beim Wort 'Birke' (deutlicher Ausschlag) dachte eine Vp wütend an einen Nachbarn, der die schönen Birken im Garten gefällt hatte.
- Die starken elektrodermalen Reaktionen auf das Wort 'Stift' erklärte eine Vp damit, daß sie als Kind mit der Aussprache dieses Wortes große Schwierigkeiten hatte.

In diesem letzten Fall wurde die Erregung nicht durch das Bezeichnete (Wortinhalt, signifié, SAUSSURE) hervorgerufen, sondern durch die Wortgestalt (Lautgestalt, signifiant, SAUSSURE).

Vorversuche der Verfasserin weisen darauf hin, daß Wörter, die Emotionen bezeichnen (Angst, Ärger, Freude, ...) fast keine elektrodermalen Reaktionen bewirken. Das Wort 'Ärger' scheint also keinen Ärger auszulösen.

2.2.3. Verhaltensmessung

Die dritte Ebene, auf der versucht wird, Emotionen zu erfassen, ist die Verhaltensebene. Zur Bestimmung der Emotionalität von Wörtern wird vor allem das Verhalten im Assoziationsversuch herangezogen. Es gibt verschiedene Indikatoren: a) die Reaktionszeit, b) die Art der Assoziation, c) Reproduktionsfehler.

2.2.3.1. Assoziationsverhalten

a) Reaktionszeit (RT)

An der Reaktionszeit versucht man abzulesen, ob es sich bei dem Stimuluswort um ein emotional geladenes Wort handelt. Eine verlängerte Reaktionszeit deutet auf ein 'emotionales' Wort hin. EPSTEIN & FENZ (1962, zit. nach CRAMER 1968) fanden, daß die Länge der Reaktionszeit zunahm, wenn

die Stimuli relevanter waren. Ihren Vpn (Fallschirmspringer) legten sie als relevante Stimuli Wörter vor, die unmittelbar mit dem Absprung zu tun hatten. Aus der Länge der Reaktionszeit ließe sich eine Intensitätsbestimmung ableiten, indem man z.B. die Wörter anhand der RT in eine Rangreihe bringt. W.SMITH (1922/1970) ordnete die 100 Wörter seiner Untersuchung nach der durchschnittlichen Reaktionszeit seiner 50 Vpn an. Laut SMITH wird die Reaktionszeit durch die Intensität des affektiven Tons unabhängig von seiner Qualität verlängert.

In anderen Untersuchungen zeigte sich aber, daß die RT nur bei unangenehmen Wörtern verlängert ist, bei angenehmen Stimuli hingegen sogar verkürzt (TOLMAN & JOHNSON 1918, zit. nach RAPAPORT 1977). BROWN (1971), der einen Überblick über verschiedene Studien bringt, kommt ebenfalls zu dieser Abfolge in der RT: A, N, U. Eine sehr kurze Reaktionszeit wäre ein Zeichen für die angenehme Qualität eines Wortes, eine lange RT für die unangenehme Qualität. Bei neutralen Wörtern ergibt sich eine mittlere RT. Die Ergebnisse sind jedoch nicht einheitlich. So fand CORDES (1901, zit. nach RAPAPORT 1977) eine Verlängerung bei A-Stimuli.

b) Art der Assoziation

Neben der Reaktionszeit ist ein weiterer Indikator die Art der Assoziation. Emotionale Wörter unterscheiden sich von neutralen Wörtern durch die größere Variabilität der Erstreaktionen über die Vpn (Maß D 'discrete association', nur eine Reaktion pro Vp). Beim fortgesetzten Assoziieren ('continued association', Maß m) ergeben sich innerhalb der Vp weniger verschiedene Reaktionen auf emotionale Wörter im Vergleich zu den neutralen Wörtern (CRAMER 1968). Affektive Wörter rufen mehr idiosynkratische Assoziationen hervor als neutrale Stimuli. Ungewöhnliche Assoziationen, Klangassoziationen usw. deuten auf eine dem Stimuluswort zugrundeliegende Emotion (oder emotionalen Komplex) hin.

Die positive oder negative emotionale Färbung eines Wortes läßt sich über den Inhalt der Assoziationen erschließen. Es zeigt sich, daß angenehme Wörter angenehme Assoziationen hervorrufen, während unangenehme Wörter unangenehme Assoziationen auslösen. (STAATS & STAATS 1959, GEER & MOLLENAUER 1964; zit. nach CRAMER 1968)

c) Reproduktionsfehler

Ein weiteres Indiz für die Emotionalität eines Wortes sind Reproduktionsfehler bei der Wiederholung des Assoziationsversuchs. Die Vp ist nicht in der Lage mit dem gleichen Wort auf den Stimulus zu reagieren wie beim ersten Mal.

Bei einer Reihe von Untersuchungen wurde nur ein Indikator herangezogen, bei manchen eine Kombination der Indikatoren. Hier gilt ein Wort erst als emotional bedeutsam, wenn es in mehreren Indikatoren Auffälligkeiten zeigt.

2.2.3.2. Probleme der Assoziationsparameter

Die assoziativen Schwierigkeiten können auch von anderen Faktoren hervorgerufen werden. So macht LAFFAL (1955, zit. nach BROWN 1971) nichtemotionale Gründe für die assoziativen Schwierigkeiten verantwortlich. Eine verlängerte Reaktionszeit und Reproduktionsfehler kommen bei Wörtern vor, die keine dominante Primärantwort besitzen und die viele gleichwertige Antwortalternativen aufweisen.

PETERSON & JUNG (1907) hielten die Erklärungsansätze, die eine lange RT durch die Ungewöhnlichkeit und Schwierigkeit des Stimuluswortes zu erklären versuchten, für nicht ausreichend, da einzelne Vpn auch auf leichte Wörter, falls sie bei ihnen eine emotionale Reaktion auslösten, eine längere RT hatten. Die Wörter 'water', 'lake', 'swim' wirkten in dieser Weise bei einer Vp, die sich ertränken wollte.

Als eine weitere mögliche Quelle für die assoziativen Schwierigkeiten wird die Abstraktheit des Stimuluswortes angesehen. M.G.SMITH & HARLESTON (1966) zweifeln die Schlüsse von

LEVINGER & CLARK (1961) an, die Effekte (längere RT und Reproduktionsfehler) der Emotionalität des Stimulus auf die Wortassoziationen aufzeigten. Ihrer Meinung nach kommen die Effekte dadurch zustande, daß die emotionalen Wörter von LEVINGER & CLARK abstrakter sind. In ihrer Untersuchung, in der sie beide Faktoren kombinieren und variieren, ergibt sich, daß nur die Abstraktheit der Wörter eine verlängerte Assoziationszeit und Reproduktionsfehler bedingt. Auch rufen nur abstrakte Wörter fast zweimal so viele verschiedene Assoziationen hervor wie konkrete Wörter. Emotionale und neutrale Wörter unterscheiden sich nicht. SMITH & HARLESTON folgern daraus, daß die Emotionalität das assoziative Verhalten nicht signifikant beeinflußt, wenn die Abstraktheit des Stimulus kontrolliert wird.

BROWN (1971) wies bei einer Überprüfung von 10 Stimuluslisten (aus 12 Wortassoziationsstudien) nach, daß in acht Fällen eine signifikant negative Korrelation zwischen der Emotionalität (E) und der Konkretheit (C) vorlag. Die emotionalen Wörter waren abstrakter als die neutralen Wörter. BROWN vermutet, daß die Konkretheit die wichtigere Variable ist. Die assoziativen Schwierigkeiten wären also auf die Abstraktheit des Stimuluswortes zurückzuführen und nicht auf die Emotionalität.

2.2.3.3. Sentogramm

Abschließend soll noch der eigenwillige Versuch von CLYNES (in LEVI 1975) erwähnt werden. CLYNES versucht, die Emotionalität von Wörtern anhand des Fingerdrucks zu untersuchen. Mit einem Sentograph mißt er die vertikalen und horizontalen Vektorkomponenten des Fingerdrucks. Dabei ist es die Aufgabe der Vp, auf ein Wort* hin, das ihr genannt wird, einen Knopf jeweils so stark niederzudrücken, daß sie dadurch genau die Emotion ausdrückt, die durch das Wort bei

* Beispielsweise: Vater, Mutter, Sohn, Nixon, Frau, Freund, Geschäftsfeind, rot, gelb, blau, grün, Liebe, Sex, Eifersucht, Freude, Haß, Gram, Ärger.

ihr angesprochen wird. Die Annahme von CLYNES u.a. ist, daß sich Gefühle in jeglichem Verhalten ausdrücken lassen. Er erhielt je nach Wort unterschiedliche Sentogramme, die er in Beziehung zur Emotionalität des genannten Wortes setzt. Weitere Verhaltensdimensionen sind bisher in der Forschung nicht berücksichtigt worden.

2.3. Übereinstimmung zwischen den verschiedenen Maßen

2.3.1. Subjektive Einschätzung und physiologische Reaktion

WECHSLER (1925) stellt fest, daß viele Autoren eine Korrelation zwischen subjektivem Urteil und EDR annehmen (z.B. VERAGUTH; RADECKI 1911), daß aber nur wenige diese Beziehung untersucht hätten, so z.B. WELLS & FORBES (1911). Sie ließen von ihren beiden Vpn Wörter nach vier Intensitätsgraden beurteilen und erhoben zusätzlich die elektrodermalen Reaktionen auf diese Wörter. Die stärker emotional beurteilten Wörter wiesen im Durchschnitt größere Ausschläge auf als die schwächer emotionalen Wörter. Die vielen Ausnahmen und die geringe Anzahl der Vpn erlauben keine verbindliche Aussage. WECHSLER berechnet in seiner Untersuchung auf der Basis der zusammengefaßten Daten von 10 Vpn für zwei Listen à 13 Wörter folgende Rangkorrelationen zwischen EDR-Amplitude und SE (5 Intensitätsgrade): .59 und .67. Er interpretiert das als relativ gute Übereinstimmung, weist aber auch auf Diskrepanzen hin: manche Wörter, die keine EDR auslösten, wurden als emotional beurteilt, andere, bei denen eine EDR vorlag, waren nach Einschätzung der Vp neutral.

SYZ (1926/27 b) erhält in seiner Untersuchung deutliche Unterschiede zwischen EDR und SE. Von 50 Wörtern erhob er bei 50 Vpn die EDR und das Urteil, ob das Wort emotionale Reaktionen hervorgerufen hat. Pro Wort registriert er die Häufigkeit von EDR und berichteter Emotion (siehe Abb. 4). Vor allem bei Wörtern aus dem sexuellen Bereich (kiss, naked body, intercourse) traten bei der Mehrzahl der Vpn EDR auf,

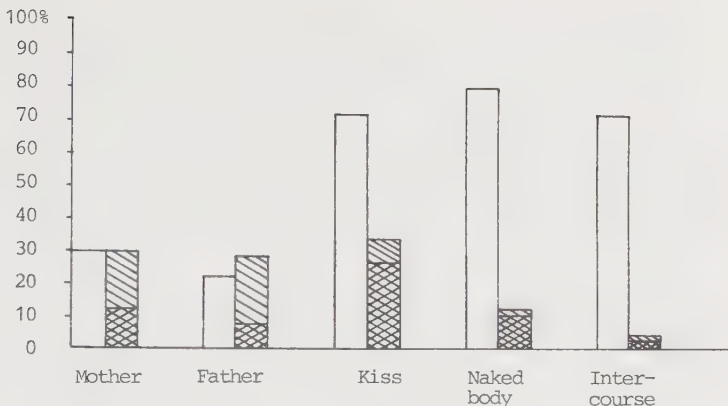


Abb. 4: Prozentuale Häufigkeit elektrodermalen Reaktionen (leere Säule) und subjektiv eingeschätzter Emotionen (schraffierte Säule).
Kreuzschraffur: Übereinstimmung zwischen EDR u. SE
Einfachschraffur: Fälschlich angegebene Emotionen.

trotzdem wurde nur vereinzelt von Emotionen berichtet. Andererseits war von Emotionen die Rede bei Wörtern wie 'mother' und 'father', obwohl ein Teil der Vpn hier keine EDR aufwiesen. SYZ zieht aus seinen Daten den Schluß, daß das subjektive Urteil unzuverlässig sei und die EDR der bessere Indikator zur Erfassung von emotionalen Reaktionen ist.

Auch die Beziehung zwischen EDR und beurteilter Qualität (A - U) wurde untersucht. DYSINGER (1931, zit. nach WOODWORTH & SCHLOSBERG 1965, S.155) ließ 150 Wörter auf einer 5-Punkte Skala mit den Polen 'sehr angenehm' und 'sehr unangenehm' einschätzen, zusätzlich zur EDR-Messung. Er erhielt folgende Ergebnisse:

Kategorie	mittlere EDR
sehr angenehm	129
angenehm	98
indifferent	79
unangenehm	101
sehr unangenehm	147

Tabelle 4: Mittlere EDR-Werte der qualitativen Urteilkategorien.

Die neutral beurteilten Wörter hatten die niedrigsten Werte; je stärker die mit den Wörtern verbundenen Gefühle waren, desto höher wurden die Werte.

LANIER (1941. zit. nach WOODWORTH & SCHLOSBERG 1965) ermittelte die höchsten EDR-Ausschläge bei der Gruppe mit 'gemischten Gefühlen'. Die anderen drei Kategorien (indifferent, angenehm und unangenehm) unterschieden sich kaum in den EDR-Werten.

MCCURDY (1950) überprüft die Beziehung zwischen EDR und berichteter Stärke der Emotion bei vielen Untersuchungen, einschließlich der oben angeführten, und erhält konsistent hohe Korrelationskoeffizienten. Aus den Daten von WELLS & FORBES (1911) berechnet er einen Korrelationskoeffizienten von .93. Anhand der Daten von SYZ gewinnt er eine Korrelation von .45.* Bzgl. DYSINGERS Daten erhält MCCURDY ein C von .87. Bei seiner eigenen Untersuchung zeigte er 51 Studentinnen 20 Adjektive verbunden mit der Frage "Are you..." und der Aufforderung mit 'ja' oder 'nein' zu antworten. Die EDR wurde während des Versuchs aufgezeichnet, anschließend brachten die Vpn die Wörter in eine Rangreihe analog zum emotionalen Wert. Die Rangkorrelation zwischen Beurteilung und EDR betrug .76.

BROWN (1971) berechnete von drei Untersuchungen die Korrelation zwischen EDR-Werten und den E- und P-Werten dieser Wörter (von BROWN & URE 1969) und erhielt folgende Korrelationskoeffizienten:

EDR-Werte von	$r_{EDR,E}$	$r_{EDR,P}$
W.SMITH (1922)	.62**	-.02
M.G.SMITH & HARLESTON (1966)	.55**	-.16
LEVINGER & CLARK (1961)	.45**	.07

Tabelle 5: Korrelationen zwischen EDR und der E- bzw. P-Skala.

* Meiner Ansicht nach erlauben die Daten von SYZ keine derartige Berechnung, da ja Syz gerade demonstriert, daß bei vielen Wörtern die Vpn, die eine EDR hatten, und die Vpn, die hier von einer emotionalen Reaktion berichten, nicht die gleichen sind.

** $p < .01$ (S.360)

Er stellt eine enge Beziehung zwischen der beurteilten emotionalen Intensität und der elektrodermalen Reaktion fest, jedoch keine Beziehung zwischen der beurteilten emotionalen Qualität (angenehm - unangenehm) und der EDR.

Eine Untersuchung im deutschen Sprachraum zur Frage des Zusammenhangs zwischen der Intensität der physiologischen Veränderung und der Intensität der erlebten Emotion stammt von TRAXEL & BECHER (1957). Sie boten ihren 19 Vpn 10 Wortpaare an, bei denen diese jeweils beurteilen sollten, welches der zwei Wörter eine stärkere 'gefühlsmäßige' Wirkung gehabt habe, und maßen gleichzeitig den Hautwiderstand. Die biseriale Korrelation zwischen Urteil und EDR-Amplitude betrug .61. Der Schätzwert der wahren Korrelation zwischen der Größe der Widerstandsänderungen und dem Urteil ergab .98. Dieser theoretische Wert von .98 wird erwartet, wenn beide Merkmale vollkommen zuverlässig sind.* Aus dieser hohen Korrelation schlossen sie, "daß man berechtigt ist, .. aus der Größe ... von Hautwiderstandsänderungen Rückschlüsse auf die Intensität der emotionalen Veränderung zu ziehen." (S.300)

Zusammenfassend kann gesagt werden, daß zwischen der Intensitätsbeurteilung und den physiologischen Daten ein Zusammenhang besteht, der aber nicht vollkommen ist.** Jedes Maß erfaßt einen etwas anderen Aspekt des emotionalen Erregungspotentials. Von einigen Forschern (SYZ 1926/27 b, TRAXEL & BECHER 1957) wird dabei der physiologischen Messung der Vorzug gegeben. Sie nehmen an, daß die physiologischen Daten eher die Affekte widerspiegeln im Gegensatz zu den subjektiven Aussagen der Vpn, die vielen Fehlermöglichkeiten unterliegen.

* Stabilitätskoeffizient der Urteile: .60 (siehe S.43)
Stabilitätskoeffizient der EDR: .64 (siehe S.65/66)

$$\frac{.61}{\sqrt{.64} \cdot \sqrt{.60}} = .98 \quad (\text{TRAXEL \& BECHER, S.295})$$

** Siehe auch S.45: Diskrepanzen zwischen dem Urteil der Vp und den physiologischen Werten bei MANNING & MELCHIORI (1974).

2.3.2. Subjektive Einschätzung und Assoziationsverhalten

Subjektive Einschätzung und Reaktionszeit

BROWN (1971) führte mehrere Korrelationsberechnungen durch. Er korrelierte die Reaktionszeitwerte von acht Assoziationsstudien mit den E- und P-Einschätzungen der Wörter (auf einer 7-Punkte Skala, BROWN & URE 1969). Bzgl. E erhielt er vier signifikante Korrelationskoeffizienten. Bei vier Studien zeigte sich kein Zusammenhang. Bzgl. P erbrachten zwei Studien eine auf dem 5%-Niveau signifikant negative Korrelation. Die anderen Koeffizienten waren nicht signifikant.

RT-Werte von	$r_{RT,E}$	$r_{RT,P}$
W.SMITH (1922)	.47**	-.11
LEVINGER & CLARK (1961)	.43**	-.27*
RAPAPORT et al. (1946)	.33**	-.20
TENDLER (1945)	.25*	-.17
SMITH & HARLESTON (1966)	.19	-.24
WALLENHORST (1965)	.16	-.15
BROWN & OGLE (1966)	.07	-.26
BROWN (1965)	.07	-.22*

Tabelle 6: Korrelationen zwischen RT und der E- bzw. P-Skala. (BROWN 1971, S.360)

Subjektive Einschätzung und Reproduktionsfehler

Bei vier Studien überprüfte BROWN (1971) den Zusammenhang zwischen Reproduktionsfehlern und den Einschätzungen auf der E- und der P-Skala (Daten von BROWN & URE 1969). Drei Studien wiesen eine auf dem 1%-Niveau signifikant positive Korrelation zwischen Reproduktionsfehlern und E auf. Zwei der Studien korrelierten signifikant negativ mit P.

Reproduktionsfehler (RF) Werte von	$r_{RF,E}$	$r_{RF,P}$
LEVINGER & CLARK (1961)	.52**	-.41**
BOWN (1965)	.30**	-.45**
TENDLER (1945)	.26**	.12
SMITH & HARLESTON (1966)	.19	-.20

Tabelle 7: Korrelationen zwischen Reproduktionsfehlern und der E- Bzw. P-Skala. (BROWN 1971, S.360)

* $p < .05$

** $p < .01$

Da bis auf die Studien von BROWN (1965) und BROWN & OGLE (1966) eine Konfundierung von E und C (Konkretheit) vorliegt (signifikant negative Korrelation zwischen E und C), lassen sich die verlängerten Reaktionszeiten und Reproduktionsfehler nicht eindeutig auf E zurückführen, sie könnten auch von der Abstraktheit dieser Wörter herrühren.

2.3.3. Physiologische Reaktion und Assoziationsverhalten

PETERSON & JUNG (1907) fanden höhere EDR-Ausschläge bei Wörtern, die eine längere Reaktionszeit aufwiesen, und bei Wörtern, die Reproduktionsfehler zeigten, basierend auf den Werten von vier Vpn.

W.SMITH (1922/1970) erhielt ähnliche Rangreihen bei der Anordnung seiner 100 Wörter nach dem mittleren EDR-Ausschlag und nach der mittleren Reaktionszeit. Als Korrelationskoeffizient der beiden Reihen gibt er einen Wert von .47 an.

W.SMITH hält die EDR für das bessere Maß, da es differenzierter mißt, über eine größere Spannbreite variiert (Verhältnis größter zu kleinstem Wert) und seiner Meinung nach gültiger ist. Bei manchen indifferenten Wörtern ergab sich eine lange Reaktionszeit, aber weder EDR noch Reproduktionsfehler. Die EDR unterliegt auch nicht der willentlichen Kontrolle wie die Reaktionszeit.

Bei EPSTEIN & FENZ (1962, zit. nach CRAMER 1968) waren die EDR-Werte konsistent mit denen der Reaktionszeit. Die Größe der EDR-Amplitude und die RT erwiesen sich als Funktion der Relevanz des Stimulus.

Zusammenfassung

Es gibt die verschiedensten Versuche, die emotionale Tönung von Wörtern zu messen: mit Hilfe von Beurteilungen, physiologischen Reaktionen und Verhaltensindikatoren. Keine dieser Methoden ist fehlerfrei, keine kann die Emotionalität voll-

kommen erfassen oder getreu abbilden. Die Reaktion, die man mißt, wird nicht nur durch die Emotionalität hervorgerufen, sondern auch durch andere Faktoren beeinflusst.

- Die subjektive Einschätzung kann von sozialen Faktoren und der Offenheit der Vp abhängen.
- Die elektrodermale Reaktion unterliegt einem Positionseffekt und ist anfällig für Störungen jeglicher Art.
- Das Assoziationsverhalten wird mitbestimmt durch weitere Wortparameter wie z.B. die Abstraktheit, Vorkommenshäufigkeit u.a..

Das Assoziationsverhalten erweist sich als ungünstigster Indikator, da es durch andere Wortcharakteristika beeinflusst wird, die sich nicht beliebig abtrennen lassen.

Die subjektive Einschätzung wird von einigen Experimentatoren abgelehnt oder als stärker fehlerbelastet angesehen als die EDR-Messung.

Die EDR schneidet als bester Indikator ab. Die Schwierigkeiten dieser Methode liegen in der Realisierung einer sorgfältigen Messung (absolut störungsfreie Versuchssituation, Beachten von Orientierungsreaktionen und Habituation).

2.4. Das emotionale Erregungspotential

Nachdem nun das Problem der emotionalen Tönung von Wörtern unter verschiedenen Gesichtspunkten erörtert wurde, soll an dieser Stelle der Versuch einer Definition des 'emotionalen Erregungspotentials' erfolgen.

Das 'emotionale Erregungspotential' gibt die Stärke der Erregung an, die ein Wort auslöst. Die Qualität der durch ein Wort hervorgerufenen Emotion bleibt dabei unberücksichtigt, es wird nur der Intensitätsaspekt des emotionalen Gehaltes angesprochen.

Man kann je nach Operationalisierung verschiedene Aspekte des emotionalen Erregungspotential erfassen: 1) die physio-

logische, vegetative Aktivierung und 2) die subjektiv empfundenen Erregung.*

Die Wörter lassen sich danach kennzeichnen, ob sie eine emotionale Erregung auslösen, und wenn ja, wie stark die ausgelöste Erregung ist.

Ob ein Wort eine emotionale Erregung hervorruft, hängt vom Individuum, seiner Lerngeschichte und der Situation ab. Ein Wort hat keine emotionale Bedeutung an sich, sondern nur für eine bestimmte Vp in einer bestimmten Zeit und in einem bestimmten Kontext.

* Die Verhaltensebene bleibt ausgeklammert, da keine befriedigenden Indikatoren existieren.

3. Wortcharakteristika und verbales Lernen

3.1. Emotionale Tönung

Im folgenden werden die Untersuchungen zum Einfluß der emotionalen Tönung von Wörtern auf das verbale Lernen dargestellt. Zunächst soll der Einfluß der emotionalen Qualität besprochen werden, anschließend wird ausführlicher auf die Wirkung der emotionalen Intensität eingegangen.

3.1.1. Qualität der emotionalen Tönung

Untersuchungen, die sich mit der Beziehung der emotionalen Tönung von Wörtern zum verbalen Lernen befaßten, wurden vor allem in den Zwanziger und Dreißiger Jahren unseres Jahrhunderts durchgeführt. Danach geriet dieser Themenbereich ziemlich in Vergessenheit und wurde erst wieder in den Sechziger Jahren aufgegriffen. (ANISFELD & LAMBERT 1966)

Ursprünglich galt das Interesse nur der Qualität der emotionalen Tönung. Man wollte feststellen, ob angenehme Wörter schneller gelernt werden als unangenehme. Einige Studien verwendeten nur 'angenehme' (A) und 'unangenehme' (U) Wörter, manche bezogen noch eine dritte Kategorie mit ein: 'neutrale' oder 'indifferente' (N) Wörter. Hier sollte geprüft werden, wie gut die neutralen Wörter im Vergleich zu den angenehmen und unangenehmen Wörtern gelernt werden. Nehmen die unangenehmen Wörter die letzte Stelle ein oder werden die neutralen Wörter am schlechtesten behalten?

Hinter der Mehrzahl der Versuche stand keine explizite Theorie. Ein Teil der Forscher bemühte sich, die Verdrängungstheorie von FREUD nachzuweisen, während andere versuchten, sie zu widerlegen. Manche Experimentatoren hatten auch nur die Vorstellung eines allgemein hedonistischen Prinzips als Motiv für ihre Untersuchungen.

TAIT (1913)* bot seinen Vpn A-, U- und N-Wörter einmal dar und prüfte anschließend die Behaltensleistung. Am häufigsten wurden die A-Wörter erinnert, dann folgten die U-Wörter. Die N-Wörter nahmen den letzten Platz ein.

TOLMAN (1917)* ließ A-, U- und N-Wörter lernen. Wieder zeigte sich eine Tendenz des Überwiegens der A-Wörter, allerdings wurden hier am zweitbesten die indifferenten Wörter erinnert und am schlechtesten die U-Wörter.

Bei den eben erwähnten Versuchen bestimmte der Experimentator die Qualität der affektiven Tönung. Man muß sich fragen, wie reliabel und valide diese Beurteilungen sind.

Die Untersuchung von CHANEY & LAUER (1929)*, in der die Qualität der Wörter durch Gruppenurteil erfaßt wurde, erbrachte keine schlüssigen Ergebnisse.

RAPAPORT (1977) findet deutlichere Ergebnisse bei Gedächtnisversuchen, bei denen die Vp selbst die Beurteilung der affektiven Qualität vornimmt (z.B. STAGNER 1933). Meist ergab sich die Abfolge A - U - N, in manchen Fällen lautete sie A - N - U. Letzteres Ergebnis wurde im Einklang mit der Verdrängungstheorie von FREUD gesehen. Ersteres wies auf eine Überlegenheit emotional getönter Wörter gegenüber neutralen hin.

THOMSON (1930) ließ ihre Vpn je 20 sehr angenehme und sehr unangenehme Wörter aufschreiben. Einen Monat später führte sie eine Behaltensprüfung durch. Die A-Wörter wurden signifikant besser erinnert.

In der Untersuchung von SHARP (1938, zit. nach McGEACH & IRION 1952), in der sie persönlich bedeutsame (akzeptierbare und unakzeptierbare) und neutrale Wörter lernen ließ und Tage später das Behalten der Wörter prüfte, zeigte sich, daß neutrale Wörter leichter gelernt werden als die persönlich bedeutsamen Wörter, daß aber später die akzeptablen am besten erinnert werden und die unakzeptablen am schlechtesten.

* zit. nach RAPAPORT 1977

Zusammenfassend kann man über die früheren Versuche sagen, daß sie zu keinem eindeutigen Ergebnis führten. In manchen Studien ergaben sich keine Unterschiede, in einigen die Abfolge A - N - U und in anderen die Abfolge A - U - N. Die Unterschiede waren teilweise sehr schwach und wurden meistens nicht auf ihre Signifikanz geprüft. Die inkonsistenten Ergebnisse sind u.a. auf methodologische Fehler zurückzuführen. Relevante Variablen, wie z.B. die Vorkommenshäufigkeit, Konkretheit und andere Wortcharakteristika wurden nicht kontrolliert. Überprüft wurden auch nicht die Richtigkeit der Qualitätsurteile und die möglichen Unterschiede in der Stärke der emotionalen Tönung.

Die Ergebnisse der neueren Untersuchungen sind gleichfalls sehr uneinheitlich. Zuerst sollen Untersuchungen angeführt werden, die nur A- und U-Wörter verglichen.

Die Experimente von ANISFELD & LAMBERT (1966) zum Free Recall erbrachten keine signifikanten Unterschiede zwischen A- und U-Wörtern. Die Lernliste bestand aus 12 angenehmen und 12 unangenehmen Wörtern, die hinsichtlich der Häufigkeit (THORNDIKE-LORGE frequency) und der Bedeutungshaltigkeit (m) parallelisiert waren. Experiment 1 und 2 prüften das intentionale Lernen. Die Wörter wurden visuell dargeboten, bei Experiment 1 betrug die Zeit 3.8 sec pro Wort, beim zweiten Experiment 1.5 sec pro Wort. Die Vpn hatten anschließend drei Minuten Zeit für die schriftliche Behaltensprüfung. Unter der Bedingung 3.8 sec wurden etwas mehr U- ($\bar{X}_U = 6.36$) als A- ($\bar{X}_A = 6.16$) Wörter erinnert, wenn die Wörter nur 1.5 sec gezeigt wurden, wurden weniger U-Wörter behalten ($\bar{X}_U = 4.9$, $\bar{X}_A = 6$). Beide Unterschiede waren aber nicht statistisch signifikant. Das gleiche galt für Experiment 3 (absichtliches Lernen, 5 Minuten Zeit zum Einprägen der Wortliste) und Experiment 4 (unabsichtliches Lernen, als Assoziationsversuch getarnt, 10 sec pro Wort).

ANISFELD & LAMBERT (1966) untersuchten auch das serielle Lernen und das Paarassoziationslernen. Wieder fanden sie, bis auf eine Ausnahme, keine Unterschiede zwischen A und U.

Die Verwendung sinnloser Silben als Stimuli beim Paarassoziationslernen hatte zur Folge, daß die Verbindungen mit den A-Wörtern besser gelernt wurden. ANISFELD & LAMBERT begründen dies damit, daß ein Item positive Konnotationen leichter erwirbt als negative. Die Hypothese, die sie aus ihren Ergebnissen ziehen, lautet: Die affektive Tönung spielt nur beim Erwerb einer neuen Bedeutung eine Rolle, hat aber sonst keinen Einfluß auf das verbale Lernen. In einer weiteren Untersuchung (ANISFELD 1970) konnte diese Hypothese nicht bestätigt werden.

Die Free Recall Untersuchung (unabsichtliches Lernen) von ROUTH & ELLIS (1975) erbrachte ebenfalls keinen Unterschied zwischen A- und U-Wörtern. Sie vermuten (S.65): "It is possible that the only real differences in recall occur for emotionally-toned words (whether good or bad) relative to neutral words which, unfortunately, were not included in the present study."

Wenden wir uns nun den Untersuchungen zu, die neben A- und U-Wörtern zusätzlich neutrale Wörter miteinbezogen.

AMSTER (1964) verwendete in ihrem Experiment vier verschiedene Wortlisten à 30 Wörter. Die Wörter (62 Substantive, mit gleichem Häufigkeitsrang von THORNDIKE-LORGE (> 50 pro Million), Wortlänge: 4 - 6 Buchstaben) wurden vorher von 40 Studentinnen nach ihrer angenehmen Qualität (pleasantness = Pl) auf einer 9-Punkte Skala mit den Polen 'pleasant' und 'unpleasant' eingestuft. Die vier Wortlisten wiesen folgenden Pl-Werte auf:

- | | | |
|---------------------|----------------------|------------------------|
| 1) neutrale Liste | $\bar{X}_{Pl} = 4.9$ | (Streubreite: 2.5 - 7) |
| 2) heterogene Liste | $\bar{X}_{Pl} = 5$ | (" : 1 - 3 und 7 - 9) |
| 3) gute Liste | $\bar{X}_{Pl} = 6.4$ | (" : 5 - 9) |
| 4) schlechte Liste | $\bar{X}_{Pl} = 3.6$ | (" : 1 - 6) |

Die einzelnen Wörter tauchten in verschiedenen Listen auf. So befanden sich 11 neutrale Wörter in Liste 1, 3 und 4, 15 gute in Liste 2 und 3, 15 schlechte in 2 und 4, 10 mittel-schlechte in 1 und 4, und 9 mittelgute in 1 und 3.

Die Vpn (126 Studentinnen) bekamen eine der vier Listen vorgelesen (alle 7 sec ein Wort) und sollten die Wörter auf der 9-Punkte Skala (Pl) einschätzen. Die Hälfte der Vpn wurde instruiert, die Wörter zu lernen, bei der anderen Hälfte sollte das inzidentelle Lernen geprüft werden. Am Ende der Wortdarbietung hatten die Vpn 7 Minuten Zeit zur freien Reproduktion.

Die Behaltensleistung bei der schlechten Wortliste war signifikant am schwächsten ($\bar{X} = 12.75$). Von der neutralen Liste wurden durchschnittlich 15.65 Wörter erinnert, von der guten Liste 15.75 und von der heterogenen Liste 16.34 Wörter (S.469). Innerhalb der guten Liste wurden signifikant mehr gute als neutrale Wörter erinnert (Verhältnis 60% zu 46%). In der schlechten Liste wurden signifikant mehr schlechte als neutrale Wörter erinnert (44% zu 39%). In der neutralen Liste wurden die mittelguten und mittelschlechten Wörter (55%/56%) signifikant häufiger erinnert als die neutralen Wörter (47%). In der heterogenen Liste wurden signifikant mehr gute als schlechte Wörter behalten (61% zu 47%). Unter der Bedingung des inzidentellen Lernens wurden neutrale Wörter besser erinnert, was im Gegensatz zu der Vermutung steht, daß der Einfluß der affektiven Qualität beim intentionalen Lernen schwächer wäre.

AMSTER schließt aus ihren Daten, daß angenehme Wörter besser erinnert werden als unangenehme und daß Wörter von beiden Extremen besser behalten werden als neutrale. Das Überwiegen der A-Wörter interpretiert sie folgendermaßen: gute Wörter lösen mehr Assoziationen aus und dies führt zu einer höheren Interitemassoziationsstärke der guten Wörter. Das gute Abschneiden der heterogenen Liste erklärt sie mit Hilfe des 'Von RESTORFF Effekt'. Die Ähnlichkeit innerhalb der Liste wäre hier am niedrigsten, und so würden die konnotativen extremen Wörter herausgehoben. Man kann das Ergebnis aber auch anders interpretieren. In der heterogenen Liste befanden sich hauptsächlich extrem beurteilte Wörter. Da die extrem eingeschätzten Wörter auch eine stärkere affektive

Intensität aufweisen*, ließe sich das Ergebnis auf die Intensität zurückführen.

In der Untersuchung von McNULTY & ISNOR (1967), in der die Vorkommenshäufigkeit, die Wortlänge und der Kontext kontrolliert waren, wurden die angenehmen und neutralen Wörter leichter gelernt als die unangenehmen Wörter.

In dem Experiment von GUERNETTE (1975) zum seriellen Lernen wurden die Listen (à 9 Wörter) mit positiven und die mit negativen Wörtern schneller gelernt als die Liste mit neutralen Wörtern.

Bei RHONE (1975) hingegen waren die A-Wörter den U- und N-Wörtern überlegen sowohl hinsichtlich der Schnelligkeit des Einprägens als auch der Menge der behaltenen Wörter.

Bemerkenswert ist, daß sich bei den Untersuchungen, die nur A- und U-Wörter verwendeten, kein Unterschied zeigte. Dagegen unterschieden sich bei solchen Studien, die zusätzlich N-Wörter enthielten, manchmal A und U ($A, N > U$, McNULTY & ISNOR 1967; $A > U, N$, RHONE 1975), manchmal aber auch A und U zusammen von den neutralen Wörtern ($A, U > N$, GUERNETTE 1975).

ROUTH & ELLIS (1975) stellen fest, daß auch in der neueren Literatur trotz vieler methodologischer Verbesserungen (Kontrolle der Vorkommenshäufigkeit, usw.) die Frage nach der Beziehung zwischen emotionaler Qualität und verbalem Lernen noch nicht gelöst wurde.

AUGUST, RYCHLAK, FELKER u.a. untersuchten das Behalten von gemochten (liked) und nichtgemochten (disliked) Wörtern. Es zeigten sich Wechselwirkungen mit dem Selbstkonzept und mit der Hautfarbe.

Bei RYCHLAK, CARLSEN & DUNNING (1974) lernten Studenten mit hohem Selbstkonzept Trigramme, die sie mochten, schneller

* Korrelation von $P_{Pol.}$ mit E: $r = .702$ (BROWN & URE 1969).
(siehe auch S.41 dieser Arbeit)

als die nichtgemochten. Bei Vpn mit niedrigem Selbstkonzept war das Verhältnis umgekehrt. AUGUST, RYCHLAK & FELKER (1975) erhielten das gleiche Ergebnis bei Kindern, die Substantive lernen sollten. AUGUST (1976) konnte nur den ersten Fall für weiße Kinder mit hohem Selbstkonzept nachweisen. Bei weißen Kindern mit niedrigem Selbstkonzept fand sich diesmal kein Unterschied zwischen den gemochten und nichtgemochten Wörtern. Bei AUGUST & FELKER (1977) wiederholte sich das Ergebnis von AUGUST (1976), für schwarze Kinder ergab sich aber, daß die Kinder mit hohem Selbstkonzept eine größere Neigung zeigten, die negativ beurteilten Wörter zu erinnern. Schwarze Kinder mit niedrigem Selbstkonzept wiesen wie ihre weißen Altersgenossen keinen Unterschied zwischen den positiven und negativen Wörtern auf.

Das Behalten aggressiver Wörter analysierten CAMP & ZIMET (1973) und GORDON (1976). CAMP & ZIMET boten Kindergartenkindern Verben dar, die als hoch oder niedrig aggressiv eingestuft worden waren. Drei Listen (à 6 Wörter) mit hoch aggressiven und drei Listen mit niedrig aggressiven Wörtern wurden erstellt. In beiden Gruppen gab es je eine Liste mit niedriger, mittlerer und hoher Vorkommenshäufigkeit. Sowohl die 'Aggressivität' als auch die Häufigkeit hatten eine signifikante Wirkung. Die mittelfrequenten Wörter wurden besser erinnert als die niedrig- und hochfrequenten Wörter. Die hoch aggressiven Wörter wurden häufiger reproduziert als die niedrig aggressiven. (Siehe Abb. 5)

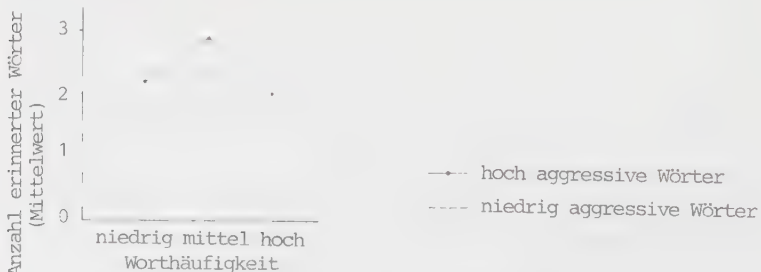


Abb. 5: Beziehung zwischen Vorkommenshäufigkeit und durchschnittlicher Behaltensleistung hoch und niedrig aggressiver Wörter. (CAMP & ZIMET, S.577)

Bei GORDON (1976) wurden von 25 stark emotional gestörten Kindern im Alter von 5 - 7 Jahren signifikant mehr niedrig aggressive Wörter erinnert. Die Häufigkeit der Wörter übte keinen Einfluß aus.

Das Behalten von Wörtern unterschiedlicher emotionaler Qualität scheint möglicherweise von der Persönlichkeit der Vp abzuhängen.*

In einem Vorversuch im SS 79 zu dieser Arbeit zeigten sich keine Unterschiede in der Behaltensleistung von positiv und negativ beurteilten Wörtern, aber Unterschiede zwischen stark und schwach erregenden Wörter. Die Vermutung, daß nicht so sehr die emotionale Qualität als vielmehr die emotionale Intensität die entscheidende Rolle beim Lernen spielt, tauchte schon in der ersten Hälfte unseres Jahrhunderts auf, u.a. bedingt durch die inkonsistenten Ergebnisse der 'Qualitäts'-Untersuchungen. In der neueren Zeit überwiegen die Untersuchungen zum Intensitätsaspekt.

3.1.2. Intensität der emotionalen Tönung

3.1.2.1. Theorien

Als Haupttheorie über die Beziehung zwischen Arousal und Gedächtnis fungiert die Konsolidierungstheorie der Michigangruppe, formuliert von WALKER (1958) und WALKER & TARTE (1963), unter Mitarbeit von KLEINSMITH und KAPLAN, ebenfalls an der Michigan-Universität. Die Theorie dieser Gruppe wird in der Literatur auch als Theorie der 'perseverativen Konsolidierung' oder als Theorie vom 'Leistungsabfall' (action decrement) bezeichnet.

Eine Konsolidierungstheorie ist zum erstenmal von MÜLLER & PILZECKER (1900, zit. nach OTT & MATTHIES 1978) aufgestellt worden. Sie sehen im Konsolidierungsprozeß die entscheidende Voraussetzung für längerfristiges Behalten. Nach

* Erinnert sei an die 'Gedächtnisoptimisten' und 'Gedächtnispessimisten' von KOWALEWSKI, S.11 und 23 dieser Arbeit.

der Reizeinwirkung dauert ein perseverativer, neuraler Prozeß an, der für die Fixierung der labilen Gedächtnisspur und damit für längeres Behalten notwendig ist. Kommt es unmittelbar nach der Reizaufnahme zu externen Störungen, kann die Information nicht stabil gespeichert werden. Für eine stabile Speicherung ist eine gewisse störungsfreie Zeitspanne nötig.

HEBB (1949, zit. nach OTT & MATTHIES 1978) machte 'reverberierende' Erregungen in Neuronennetzen für die Konsolidierung verantwortlich.

PARE (1961, zit. nach KLEINSMITH & KAPLAN 1963) greift auf die 'reverberierenden' neuralen Kreisläufe von HEBB zurück, um die Wirkung des Arousal auf die Konsolidierung zu erklären. Bei schwacher Erregung ist relativ wenig unspezifische neurale Aktivität vorhanden, um die 'reverberierende' Spur zu versorgen. Die Folge ist eine geringe Konsolidierung und schlechtes Langzeitbehalten. Eine hohe Erregung hingegen (mehr unspezifische neurale Aktivität steht zur Verfügung) führt zu mehr 'Reverberierung' und so zu einem besseren Behalten.

KLEINSMITH & KAPLAN (1963, zusammen mit TARTE) vermuten, daß die Gedächtnisspur während der Konsolidierung (Perseveration) dem Organismus nicht verfügbar ist. Die endgültige Formulierung der Theorie findet sich bei WALKER & TARTE (1963): Während der Konsolidierungsperiode, in der das bleibende Gedächtnis niedergelegt wird, besteht eine Abrufhemmung*, um die Gedächtnisspur vor Störung zu schützen. Hohe Erregung bewirkt einen intensiveren Spurenprozeß, dadurch eine größere zeitliche Abrufhemmung, aber ein besseres Langzeitgedächtnis. Vorhergesagt wird also, daß bei hohem Arousal beim Lernen die Reproduktionsleistung unmittelbar nach der Darbietung schlechter ist, später aber besser (Reminiszenzeffekt). Bei schwacher Erregung ist die Kurzzeitgedächtnisleistung

* Es werden keine Angaben gemacht, wie lange die Abrufhemmung besteht. - AMELANG, WENDT & FRÜNDT (1977) folgern aus ihrer Untersuchung, daß die Konsolidierung bei hoher Erregung (Introvertierte) intensiver verläuft und schneller abgeschlossen ist (ca. 10 sec) als bei niedriger Aktivierung (Extravertierte).

besser, die Langzeitgedächtnisleistung aber schlechter. Es liegt eine Wechselwirkung zwischen Arousal und Behaltensintervall vor.

Es werden alternative Erklärungen für die Interferenz der Erregung mit dem Kurzzeitgedächtnis (ST = short term memory) angeführt. VEHLING (1972) erwähnt zwei mögliche andere Gründe. Die Erregung wirkt störend und erhöht deshalb die Verarbeitungszeit. Ist genügend Zeit vorhanden, kommt es nicht zur Störung (COPPESEN 1969). Eine andere Interpretation wäre, daß die erregenden Items wegen der ähnlichen autonomen Reaktion einander ähnlicher wären, was das Lernen von Paarassoziationen verschlechtern würde.

Daß bei der Langzeitgedächtnisprüfung (LT = long term memory) die stark aktivierenden Items besser erinnert werden, kann daran liegen, daß die hohe Erregung zu einer verstärkten Wiederholung (Rehearsal) im Retentionsintervall führt.

Eine andere Theorie*, die Leistung und Erregungsgrad in Beziehung setzt, ist die der 'umgekehrten U-Funktion' (SCHLOSBERG 1954, HEBB 1955; zit. nach HALL 1971). Sie postuliert, daß ein mittlerer Grad an Erregung für die Leistung optimal ist. Zu geringe oder zu starke Erregung wirkt sich leistungsmindernd aus. SCHÖNPLUG (1966) bezeichnet diese Theorie als 'Theorie von der kurvenlinearen Abhängigkeit des Einprägens von der Aktivierung' und beschreibt sie folgendermaßen:

Es werden zwei gegensätzliche Wirkungen der Aktivierung angenommen, die je nach Höhe der Aktivierung unterschiedlich stark in Erscheinung treten:

* Die Untersuchungen, die hierzu durchgeführt wurden, fallen alle in den Bereich des 'Subject-Arousal', also der Erregung der Vp, der unter 4.3. behandelt wird. Hinsichtlich des Erregungspotentials des Lernmaterials fand sich keine explizite Festung. Eine Prüfung konzentriert daran, daß in den meisten Untersuchungen die Items nur in zwei Kategorien eingeteilt wurden, in stark und in schwach erregende Items. Mit zwei Punkten auf der Arousaldimension läßt sich keine kurvenlineare Beziehung nachweisen, mindestens drei Stufen wären notwendig. Ein Problem ist sicher auch, daß Würzner nicht über die ganze Erregungsdimension hin variieren, sondern nur in einem enger begrenzten Bereich bleiben.

- 1) eine die Spurenbildung fördernde Wirkung: Die Aktivierung setzt Energie frei, die die Spurenbildung beschleunigt und stärkt.
- 2) eine schädliche Wirkung: Die freigesetzte Energie beeinträchtigt die Steuerung der Vorgänge, die zur Spurenbildung führen.

Die fördernde Wirkung wächst mit der Höhe der Aktivierung negativ akzeleriert an. Die schädliche Wirkung nimmt mit der Höhe der Aktivierung positiv akzeleriert zu. Von einem bestimmten Aktivierungsniveau an übertrifft die schädliche Wirkung die fördernde.

LEVONIAN (1972) entwickelt aufgrund der unterschiedlichen Ergebnisse von 7 verschiedenen Untersuchungen* eine eigene Hypothese, die 'retentivity - accessibility' Hypothese. Er unterscheidet die Speicherstärke (retentivity) von der Abrufstärke (accessibility) und postuliert für jede eine andere Beziehung zum Arousal. Die Speicherstärke ist direkt mit dem Arousal verbunden, wächst bei allen Graden der Erregung für eine Minute an und fällt dann ab. Zwischen der Abrufstärke und der erregungsbeeinflussten Speicherstärke liegt bei Kurzzeitintervallen eine umgekehrte U-Funktion vor, bei Langzeitintervallen eine positiv lineare Funktion. Die Abrufstärke ist bei Kurzzeitintervallen bei einem mittleren Erregungsgrad am größten, bei Langzeitintervallen um so größer je höher die Erregung bei der Einspeicherung war. (Siehe Abb. 6 und 7)

LEVONIAN glaubt, mit seiner Hypothese die diskrepanten Ergebnisse erklären zu können. Bei den Untersuchungen, die das von WALKER postulierte Ergebnis** erbrachten, handelte es sich seiner Meinung nach bei den 'high-arousal' Items tatsächlich um solche mit hoher Erregung. Die Untersuchungen,

* Fünf Untersuchungen wiesen ein Ergebnis entsprechend der Theorie von WALKER auf. Zwei erbrachten auch bei der Kurzzeitgedächtnisprüfung eine bessere Leistung bei erregenden Items.

** Bei ST: schlechte Retention der 'high-arousal' Items; bei LT: gute Retention.

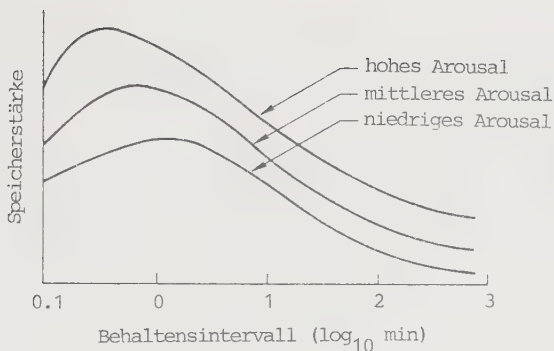


Abb. 6: Angenommene Beziehung zwischen Arousal während des Lernens und der Speicherstärke über die Zeit.*

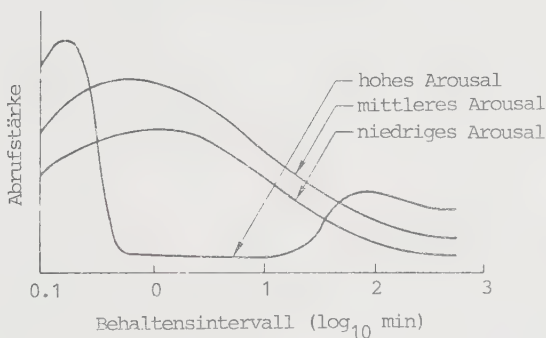


Abb. 7: Angenommene Beziehung zwischen Arousal während des Lernens und der Abrufstärke über die Zeit.*

bei denen bei ST die 'high-arousal' Items besser abschnitten, haben seiner Ansicht nach nur mittlere Erregungsgrade bewirkt. Er versucht seine Annahme nachzuweisen, daß bei den beiden Studien mit dem letzteren Ergebnis die 'stark aktivierenden' Items nur mittel erregend gewesen wären. Bei diesen beiden Studien lag eine a priori Kategorisierung der Wörter vor, im Gegensatz zu den anderen fünf Untersuchungen, die die Items a posteriori zuwiesen. Laut LEVONIAN ist bei der a posteriori Kategorisierung ein größerer Unterschied im durch-

* LEVONIAN 1972, S.302.

schnittlichen Arousal zwischen hoch und niedrig erregenden Items vorhanden als bei einer normativen oder a priori Kategorisierung.

Da eine Standardisierung der Arousal-Messung fehlt, können die Arousalwerte der verschiedenen Studien nicht direkt miteinander verglichen werden. LEVONIAN gibt keine Kriterien an, ab wann eine Erregung als hoch, mittel oder niedrig anzusehen ist. Seine Annahme, daß bei den beiden Studien nur eine mittlere Erregung vorlag, ist teilweise eine reine Spekulation.

Die 'perceived intensity of affect' Hypothese von DUTTA & KANUNGO (1975) wurde bereits auf S.24 dieser Arbeit beschrieben. Sie besagt, daß wir ein Item um so besser im Gedächtnis behalten, je stärker seine emotionale Tönung ist. Sowohl für ST als auch LT wird eine positiv lineare Beziehung zwischen der wahrgenommenen Gefühlsintensität und der Behaltensleistung angenommen.*

3.1.2.2. Untersuchungen

In den Untersuchungen zum verbalen Lernen wird die durch ein Wort hervorgerufene Erregung entweder a priori oder a posteriori festgelegt. Bei der a priori Feststellung steht vor dem Gedächtnisversuch fest, welche Wörter erregend sind, bei der a posteriori Bestimmung erfolgt die Klassifizierung der Wörter nach der im Gedächtnisversuch ermittelten Erregung. Gemessen wird die Erregung mit Hilfe der EDR oder mit Hilfe von Einschätzungen. Die Analyse der Daten geschieht entweder ipsativ oder normativ.

Die erste Gedächtnisuntersuchung, in der die Messung der EDR verwendet wurde, führte W.SMITH (1922) durch. Er maß die Ge-

* In ihren Versuchen bringen sie einen Nachweis für ihre Hypothese. Demnach werden Adjektive, die intensiver beurteilt werden, besser behalten.

fühlsbetonung von 100 Wörtern mit Hilfe der EDR in einem Assoziationstest. Von diesen 100 Wörtern wählte er 30 aus, die 41 Vpn zwei Tage nach dem Assoziationstest auswendiglernen mußten, bis sie eine fehlerfreie Reproduktion schafften. Am 5., 9., 14., 21. und 31. Tag nach dem Experiment sollten die Vpn auf eine Postkarte die Wörter aufschreiben, an die sie sich noch erinnerten und die Wörter unterstreichen, bei denen sie Schwierigkeiten hatten. 2 Punkte erhielt ein Wort, das ohne Schwierigkeiten erinnert wurde, 1 Punkt ein erinnertes unterstrichenes Wort und 0 Punkte gab es für Nichterinnern. Über die 5 Behaltensprüfungen konnte ein Wort die maximale Punktzahl von 10 erhalten (Memory Value). Die Ergebnisse des Versuchs finden sich in Abb. 8.

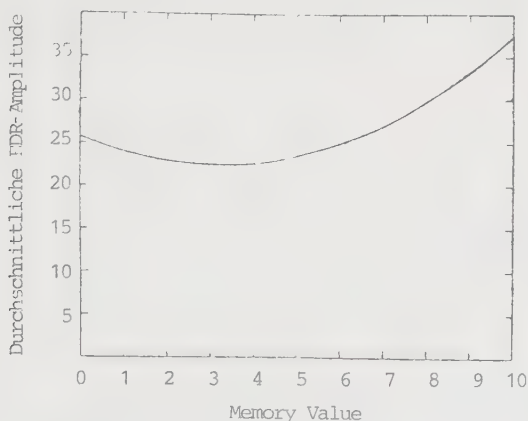


Abb. 8: Beziehung zwischen Gedächtnisleistung und EDR-Amplitude. (W.SMITH 1970, S.42)

Die am schlechtesten und die am besten erinnerten Wörter weisen eine höhere EDR-Amplitude auf. Die Wörter mit einem mittleren Gedächtniswert haben eine niedrigere EDR-Amplitude. W.SMITH vermutet, daß der Effekt noch deutlicher zutage getreten wäre, wenn die neutralen Wörter, die nicht erinnert wurden, ausgeschlossen worden wären. Dadurch hätte die Kurve bei kleinen Gedächtniswerten höhere EDR-Amplituden erreicht. Die schlecht erinnerten Wörter zeigten auch beim

Assoziationsversuch prozentual mehr Reproduktionsfehler. W.SMITH schließt, daß der affektive Ton laut EDR entweder die Erinnerung erleichtert oder hemmt. Den affektiven Ton, der die Erinnerung erleichtert, nennt er 'positive tone', den, der die Erinnerung hemmt, bezeichnet er als 'negative tone'. Diese Kategorisierung kann aber erst a posteriori durchgeführt werden, sie hat deshalb wenig Aussagekraft. In seinen theoretischen Äußerungen über den affektiven Ton definiert er den negativen Ton als auf einen Konflikt bezogen. Für den positiven Ton sei Harmonie und Verstärkung verantwortlich.

JONES (1929)* wiederholte den Versuch von SMITH und kam zu gleichen Ergebnissen. BALKEN (1933)* und STAGNER (1933)* erhielten nur Zufallskorrelationen zwischen EDR und Gedächtnisleistung.

Die neueren Untersuchungen zu diesem Aspekt laufen fast alle unter der Bezeichnung des 'Arousal'. Überblicke über Untersuchungen, die die Rolle des Arousal beim verbalen Lernen zum Thema haben, finden sich bei UEHLING (1972), LEVONIAN (1972) und M.W.EYSENCK (1976 und 1977).

Der Großteil der Versuche zur Konsolidierungstheorie von WALKER & TARTE (1963) verwendete die Paarassoziationslernmethode.

Paarassoziationsuntersuchungen

Die ersten Untersuchungen von KLEINSMITH & KAPLAN (1963) und WALKER & TARTE (1963) bestätigten die Konsolidierungstheorie. Es handelte sich um unabsichtliche Lernversuche, bei denen den Vpn acht Paare einmal dargeboten wurden. Als Stimuli dienten acht Wörter** und als Response Ziffern von 2 bis 9. Alle vier Sekunden wurde ein Dia gezeigt und zwar in folgender Reihenfolge: zuerst der Stimulus allein, dann Stimulus - Response zusammen, dann folgten zwei Dias mit je vier Farben, die die Vp zu benennen hatte. Während des Versuchs wurde die

* zit. nach RAPAPORT 1977.

** Wörter bei KLEINSMITH & KAPLAN: kiss, love, dance, money, swim, rape, vomit, exam.

EDR aufgezeichnet. Je acht Vpn nahmen an einem Versuch teil, wobei jede Vp die Paare in einer anderen Reihenfolge erhielt.* Von den insgesamt 48 Studenten (bei KLEINSMITH & KAPLAN) wurden je acht nach 2 Minuten, 20 Minuten, 45 Minuten, einem Tag und einer Woche geprüft. Der Stimulus wurde dargeboten, und die Vp hatte 8 sec Zeit, die dazugehörige Ziffer zu benennen. Pro Vp wurden die drei Wörter mit den höchsten EDR-Werten und die drei mit den niedrigsten EDR-Werten bestimmt. Gemessen wurde die Hautwiderstandsänderung in Prozent (d/R) innerhalb von 4 sec nach der Präsentation des Wortes (bei den späteren Untersuchungen erweiterten sie das Intervall auf 12 sec). Die Ergebnisse (Abb. 9) zeigen einen deutlichen Interaktionseffekt zwischen Arousal und Retentionsintervall. Schwach erregende Items führten mit der Zeit zu einem beträchtlichen Vergessen, wohingegen stark erregende Items einen Reminiszenzeffekt aufwiesen.

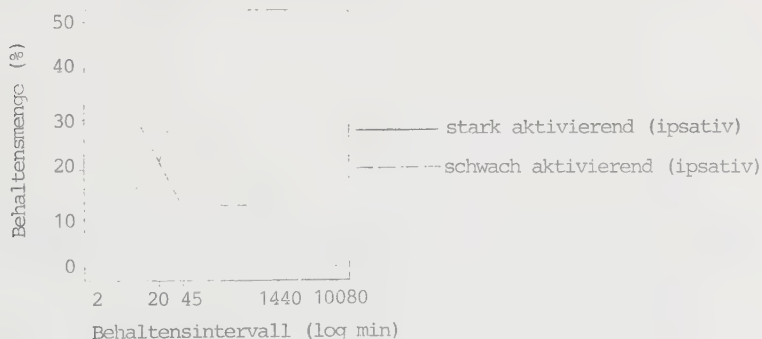


Abb. 9: Unterschiedliches Behalten von Paarassoziationen als Funktion des Arousalgrades und der Zeit. (KLEINSMITH & KAPLAN 1963, S.191)

* Damit wollen KLEINSMITH & KAPLAN den Positionseffekt kontrollieren. Da sie die Daten aber ipsativ auswerten, können sie auf diese Weise den Positionseffekt nicht ausgleichen. Die Permutation der Reihenfolge führt nur zu einer Kontrolle des Positionseffekt über die Vpn, nicht aber innerhalb der Vp. Zur Kontrolle des Positionseffekt innerhalb der Vp wären Übungsitems nötig gewesen.

WALKER & TARTE verwendeten drei verschiedene Wortlisten:

1. 'high-arousal' Liste (kiss, money, rape, vomit, slut, embrace, passion, sex)
2. 'low-arousal' Liste (walk, carrot, berry, white, glass, flower, pond, pencil)*
3. 'mixed' Liste, zusammengesetzt aus vier 'high-arousal' und vier 'low-arousal' Items.

Die Klassifikation der Wörter erfolgte a priori. Die Behaltensleistung wurde entweder 2 Minuten, 45 Minuten oder eine Woche später geprüft. Acht Studentinnen standen pro Gruppe zur Verfügung. Die Ergebnisse sind in Abb. 10 zusammengefaßt.

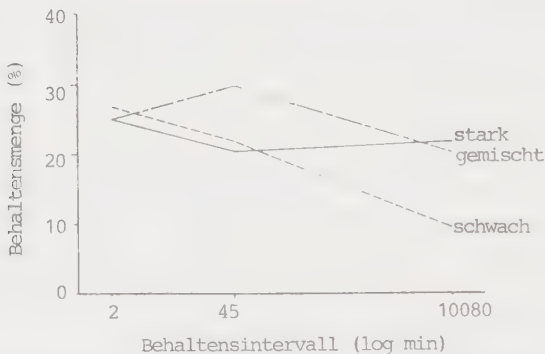


Abb. 10: Behaltensleistung bei stark und schwach erregender und gemischter Liste über die Zeit. (WALKER & TARTE 1963, S.115)

Der Effekt ist bei der unmittelbaren Reproduktion nicht so deutlich wie bei KLEINSMITH & KAPLAN, weil die Werte unabhängiger Gruppen vorliegen; bei KLEINSMITH & KAPLAN handelt es sich um abhängige Werte.

WALKER & TARTE führten auch eine Analyse ihrer Daten nach der Methode von KLEINSMITH & KAPLAN durch, bei der das Erregungspotential der Wörter anhand der individuellen EDR-Ausschläge festgestellt wird. (Siehe Abb. 11) Zahlen, die mit niedrig aktivierenden Wörtern verbunden sind, werden schneller vergessen. Ziffern, die mit stark aktivierenden Wörtern

* Diese Wörter stammen aus der Liste von W.SMITH 1922.

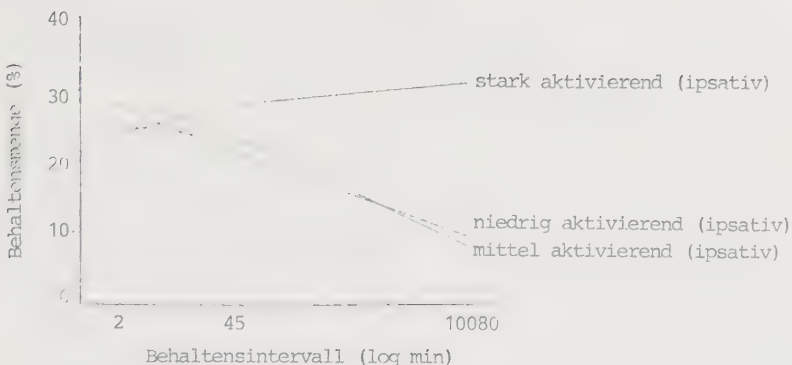


Abb. 11: Behaltensleistung bei stark, mittel und niedrig aktivierenden Wörtern (ipsativ) über die Zeit. (WALKER & TARTE 1963, S.116)

gekoppelt wurden, sind 45 Minuten und eine Woche später denen mit niedrig aktivierenden Wörtern überlegen. Die Unterschiede zwischen starker und schwacher Erregung sind signifikant.

WALKER & TARTE gehen auf den Positionseffekt ein. Bei der 2 Min. Behaltensprüfung machte sich ein starker Recency-Effekt bemerkbar, wogegen eine Woche später ein starker Primacy-Effekt auftrat. Bei der EDR-Kurve zeigten sich anfangs sehr hohe Ausschläge, die mit der Zeit abnahmen. WALKER & TARTE weisen auf die Übereinstimmung der positionsbedingten EDR-Werte mit der Reproduktionsleistung bei der Langzeitgedächtnisprüfung hin.

1964 wiederholten KLEINSMITH & KAPLAN den Versuch mit sechs Paaren, wobei diesmal als Stimuli sinnlose Silben verwendet wurden (CEF, QAP, TOV, JEX, LAJ, DAX; Assoziationswert = 0). Sie wollten das Phänomen der perseverativen Konsolidierung unabhängig von den Assoziationswerten der Wörter untersuchen. Die Behaltensleistung von je 12 Vpn wurde entweder 2 Minuten nach der Darbietung, 20 Minuten oder eine Woche später geprüft. Die Paare wurden wie zuvor anhand der individuellen EDR-Ausschläge pro Vp in drei stark erregende und drei

schwach erregende Items aufgeteilt. Wieder ergab sich eine ähnliche Kurve. Die Behaltensleistung bzgl. der niedrig aktivierenden Items war anfangs gut, später schlecht. Bei den stark aktivierenden Items war es umgekehrt.

Die Autoren versuchen zu ermitteln, ob eine mögliche Interaktion zwischen Position und Arousal besteht, da die EDR dazu tendiert zu habituierten. Sie überprüfen diesen möglichen Effekt, indem sie das 1. Item, das fast immer stark aktivierend ist, aus ihrer Analyse ausschließen. Weil sich dadurch die Ergebnisse nicht sehr verändern, schließen sie, daß keine Interaktion zwischen Position und Arousal vorliege. Diese Art der Analyse rechtfertigt keinen derartigen Schluß. Die Erregungen, die auf die sinnlosen Silben auftraten, sind alleine von Orientierungsreaktionen verursacht und nicht von der 'erregenden Qualität' der Silben. Hat sich die Vp an die Situation gewöhnt, lösen sinnlose Silben keine EDR aus (Vorversuche der Verfasserin).

In der Literatur finden sich noch weitere Untersuchungen zur Unterstützung der Konsolidierungstheorie von WALKER. KAPLAN & KAPLAN (1969) führten den Versuch* mit 6 Paarassoziationen (als Stimuli dienten die Wörter: kiss, rape, vomit, exam, money, swim) mit 189 Studenten, die möglichst entspannt sein sollten, durch. Das Gedächtnis wurde unmittelbar, 6 Minuten, 18 Minuten oder zwei Tage später geprüft. Bei der unmittelbar anschließenden Gedächtnisprüfung zeigte sich wieder die Überlegenheit der schwach erregenden Items. Nach 6 Minuten trat der Reminiszenzeffekt ein. Die stark erregenden Itemverbindungen wurden besser erinnert. Mit der Länge des Retentionsintervalls nahm die Reproduktionleistung sowohl bzgl. der stark erregenden als auch der schwach erregenden Items ab. Die Interaktion 'Arousal-Retentionsintervall' fiel schwächer aus. Die Autoren führen das auf die Entspannung zurück,

* Veränderungen gegenüber den früheren Versuchen: Zu Beginn des Versuchs ein 'Ready'-Dia. Das Dia mit dem Stimulus alleine wurde nicht mehr dargeboten.

dadurch wären die Vpn nicht so leistungsmotiviert gewesen.
(Ergebnisse siehe Abb. 12)

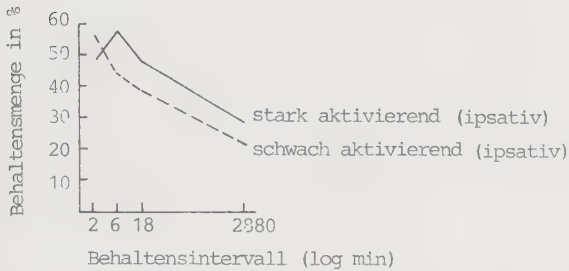


Abb. 12: Behalten von Paarassoziationen als Funktion des Arousalgrades und der Zeit. (KAPLAN & KAPLAN 1969, S.85)

CORTEEN (1969) kritisiert an den Versuchen den Einschub von Farblendias, die seiner Meinung nach zu einer Konfusion der Reaktionen führen. KAPLAN & KAPLAN (1969) geben an, daß die nach Ende der Versuche befragten Vpn vermuteten, daß ihre Reaktionen auf die Farben gemessen würden, um Aufschlüsse über ihre Persönlichkeit zu erhalten.

SCHÖNPFLUG (1966), der eine Replikation des Versuchs von KLEINSMITH & KAPLAN (1963) unternahm, erfuhr ebenfalls in einer anschließenden Befragung seiner 56 Vpn, daß diese sich vor allem auf die Farben konzentriert hatten, da sie diese als anziehender empfunden und weniger Zeit zu ihrer Benennung gehabt hatten.

SCHÖNPFLUG will mit seinem Experiment durch das zusätzliche Miteinbeziehen eines Schallpegels (Ton von 1000 Hz und 55 Phon) eine Entscheidung zwischen der Kurvilinearitätstheorie und der Theorie von WALKER herbeiführen. Nach WALKER wäre zu erwarten, daß durch die zusätzliche Aktivierung die Wiedergabe des Lernstoffs bei kurzfristigem Behalten schlechter wäre, nach längerem Behalten jedoch besser. Nach der Kurvilinearitätstheorie sollte durch eine geeignete

akustische Zusatzreizung während des Einprägens die Wiedergabe aller Verbindungen in allen Prüfungen verbessert werden. Als stark ich-bezogene Wörter verwendete SCHÖNPFLUG 'Bombe, Examen, Kuß, Nazi' und als schwach ich-bezogen 'Blume, Gedanke, Salz, Würfel'. Die eine Hälfte der Vpn erhielt während der Darbietung die akustische Zusatzreizung. Die Behaltensprüfung fand entweder zwei Minuten oder 45 Minuten nach dem Versuch statt. Anhand der aufgezeichneten Hautwiderstandskurve wurden Amplituden- und Basiswerte berechnet. Die akustische Zusatzreizung führte zu einer Erhöhung der Basiswerte. Ebenso waren die Basiswerte bei der Behaltensprüfung erhöht. Die a priori Klassifikation der Wörter ließ sich durch die mittleren Rangplätze der EDR-Amplitude bestätigen. Nach zwei Minuten wurden die Paarassoziationen besser wiedergegeben als nach 45 Minuten. Paare mit hoch ich-bezogenen Wörtern als Stimuli wurden bei allen Prüfungen besser behalten. Die akustische Zusatzreizung übte keinen signifikanten Einfluß auf die Reproduktionsleistung aus. (Ergebnisse siehe Tabelle 8)

<u>Zusatzreizung:</u> _ _ _ _ _	<u>nicht vorhanden</u>	<u>vorhanden</u>		
<u>Behaltensdauer in Minuten:</u>	2	45	2	45
Valenz des Lernstoffs				
niedrig	1.14	0.50	0.50	0.43
hoch	2.07	1.21	1.85	1.35
insgesamt	3.21	1.71	2.35	1.78

Tabelle 8: Häufigkeit richtiger Wiedergaben (Mittelwerte)
(SCHÖNPFLUG 1966, S.144)

SCHÖNPFLUG schreibt (S.145): Die Ergebnisse "stehen in deutlichem Gegensatz zu den Befunden von KLEINSMITH & KAPLAN (1963)". Er stellt die Gültigkeit der Theorie von WALKER in Frage und spricht sich für die Kurvilinearitätstheorie aus.

KAPLAN & KAPLAN (1970) glauben die Ursache gefunden zu haben für die unterschiedlichen Ergebnisse, daß einerseits Verbindungen mit niedrig aktivierenden Items bei ST besser behalten

werden als die mit stark aktivierenden und andererseits die Verbindungen mit stark aktivierenden Items auch bei ST den mit niedrig aktivierenden Items überlegen sind. Der Grund liege darin, daß bei den Untersuchungen mit dem letzteren Ergebnis die Kategorisierung nach stark und schwach erregend nicht auf ipsativen Werten beruhte, sondern auf normativen Werten (entweder a priori oder anhand der durchschnittlichen EDR-Werte über alle Vpn). Bei einer ipsativen Auswertung käme das von WALKER postulierte Ergebnis zustande, bei der normativen Auswertung zeige sich eine bessere Wiedergabe der stark erregenden Items über alle Retentionsintervalle. Eine Reanalyse* von drei Studien (KLEINSMITH & KAPLAN 1963, 1964; KAPLAN & KAPLAN 1969) bestätigte ihre Vermutung für die beiden Studien mit bedeutungshaltigem Material (siehe Abb. 13).**

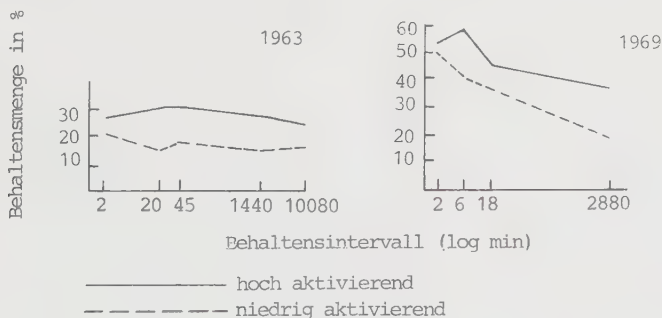


Abb. 13: Behalten von Paarassoziationen als Funktion normativer EDR-Werte und der Zeit (KAPLAN & KAPLAN 1970, S.116) (Vergleiche mit Abb.9 und 12)

* Die Items wurden nach ihrem mittleren EDR-Ausschlag in eine Rangreihe gebracht. Über dem Median befindliche Items galten als stark erregend, darunter befindliche als schwach erregend.

normativ stark erregend: rape, kiss, vomit, swim (1963)
 rape, kiss, vomit (1969)
 CEF, QAP, DAX (1964)

normativ schwach erregend: exam, love, money, dance (1963)
 exam, swim, money (1969)
 JEX, LAJ, TOV (1964)

** Die Studie mit sinnlosen Silben zeigte keine Veränderung.

KAPLAN & KAPLAN formulieren deshalb zusätzlich eine 'normative encoding' Hypothese. Bedeutungshaltiges Material wird auf eine bestimmte charakteristische Art encodiert, die in den normativen Werten widergespiegelt wird. Diese Encodierung hat einen entscheidenden Einfluß auf die Wiedergabe, unabhängig von den individuellen EDR-Werten der Vp.

BUTTER (1970) erhielt in ihrer Replikation des Versuchs von KLEINSMITH & KAPLAN (1963) die gleichen Ergebnisse wie diese. POWERS (1974) fand dagegen keine Unterschiede zwischen stark, mittel und schwach erregend beurteilten Stimuluswörtern.

Ein Hauptkritikpunkt bei den eben beschriebenen Versuchen ist, daß sie keine Übungssitems verwendeten. Darum fand eine Habituation an die Situation nicht vor, sondern erst während des Versuchs statt.

Vermutlich sind die Ergebnisse positionsbedingte Artefakte. Bei der sofortigen Behaltensprüfung zeigt sich unter der Bedingung des inzidentellen Lernens ein starker Recencyeffekt, die letzten Items werden besser reproduziert. Da bei den letzten Items wegen der Habituation eine niedrigere Erregung vorliegt, ergibt sich zwangsläufig dieser Zusammenhang zwischen niedrig erregenden Items und Kurzzeitgedächtnisleistung.

Die aufgetretenen elektrodermalen Reaktionen wurden weniger durch die Bedeutung der spezifischen Items verursacht, sondern eher durch die Gesamtsituation (rasch wechselnde Reize). Die Erregungen können daher m.E. nicht direkt auf die verwendeten Wörter bezogen werden.

MALTZMAN, KANTOR & LANGDON (1966) lehnen die Paarassoziationslernmethode ab. Nach ihrer Ansicht kommt es zu einer Konfundierung der Arousaleffekte während des Lernens und des Behaltens, da die Stimuluswörter als Hinweisreize auch bei der Behaltensprüfung dargeboten werden.

Zum Nachweis der perseverativen Konsolidierungstheorie scheint die gewählte Vorgehensweise wenig geeignet. Das Paarassozia-

tionslernen prüft, ob eine Verbindung (hier Wörter mit Zahlen von 1 bis 9) gemerkt worden ist, im Fall der beschriebenen Versuche, inwieweit die 'erregende Qualität' des Stimulus hierfür förderlich ist. Welchen Einfluß die Erregung eines Reizes auf seine Retention hat, kann damit nicht untersucht werden. Dazu sind Free Recall Untersuchungen nötig.

Free Recall Untersuchungen

Im deutschen Sprachraum fand sich in neuerer Zeit nur eine Untersuchung (von SCHÖNPFLUG & BEIKE, 1964). SCHÖNPFLUG & BEIKE untersuchten absichtliches und uabsichtliches Lernen von 12 Wörtern hoher Ich-Bezogenheit* oder 12 Wörtern niedriger Ich-Bezogenheit**. Ihre Hypothesen lauteten: Beim absichtlichen Lernen kommt es zu einer stärkeren Aktivierung, was zu einem besseren Behalten des Lernstoffs führt. Ein Stoff mit höherem Ich-Bezug erfährt stärkere Zuwendung, durch ihn wird mehr Energie freigesetzt, die dem Einprägen zufließt. Erwartet wird eine bessere Lernleistung beim absichtlichen Lernen gegenüber dem unabsichtlichen und beim persönlich bedeutsamen Material gegenüber dem mit geringerem Ich-Bezug. Das Experiment wurde mit 120 Studenten und Angestellten der Frankfurter Universität durchgeführt. Zur Gewöhnung wurden fünf farbige Dias mit Plakatabbildungen gezeigt.*** Alle fünf Sekunden erfolgte ein Diawechsel. Im Anschluß an die Darbietung hatten die Vpn 45 Sekunden Zeit für die freie Reproduktion. Während des Versuchs wurde der Hautwiderstand aufgezeichnet. Basiswerte und Amplitudenwerte wurden nach der Methode von SCHÖNPFLUG (1963) berechnet, jeweils für ein Intervall von 20 Sekunden (entspricht vier Dias). Die Varianzanalyse ergab, daß die Basiswerte und die Amplitudensummenwerte mit der Absichtlichkeit des Lernens und mit der

* Angst, Bombe, Erblindung, Erhängen, Examen, Heirat, Krieg, Kuß, Liebe, Nazi, Tanz, Überfall.

** Baum, Beleidigung, Blume, Buch, Gedanke, Gras, Gruß, Salz, Schnupfen, Tee, Wetter, Würfel.

*** Diese Dias gewöhnen nur an die Situation, aber nicht an die Stimuli. Bei den ersten Wörtern treten deshalb Orientierungsreaktionen auf.

Ich-Bezogenheit anstiegen. Außerdem fielen die Amplitudensummenwerte im Verlauf der Darbietung ab.* (siehe Abb. 14)
Die Anzahl richtiger Wiedergaben nahm sowohl mit der Absichtlichkeit des Lernens als auch mit dem Ich-Bezug des Lernstoffs zu (siehe Abb. 15).

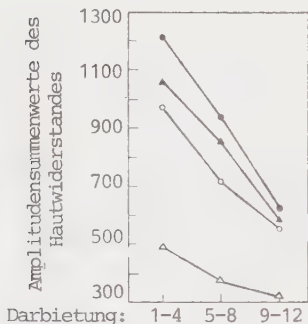


Abb. 14: Amplitudensummenwerte bei unterschiedlicher Absichtlichkeit des Lernens und verschieden hoher Ich-Bezogenheit des Lernstoffs. (SCHÖPFLUG & BEIKE, S.373)

Absichtlichkeit des Lernens

- hoch
- hoch
- ▲ niedrig
- △ niedrig

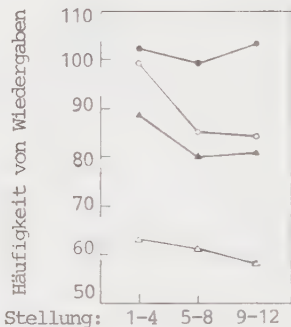


Abb. 15: Häufigkeit richtiger Wiedergaben bei unterschiedlicher Absichtlichkeit des Lernens und verschieden hoher Ich-Bezogenheit des Lernstoffs. (SCHÖPFLUG & BEIKE, S.370)

Ich-Bezogenheit des Lernstoffs

- hoch
- niedrig
- hoch
- niedrig

Die Vpn erklärten bei einer anschließenden Befragung, daß sie versucht hätten, die Wörter entsprechend ihrer Beziehungen untereinander zusammenzufassen. Der Effekt war bei den ich-bezogenen Wörtern stärker.** SCHÖNPFLUG & BEIKE ziehen aus ihrer Untersuchung den Schluß, daß Aktivierungsvorgänge für das Behalten von Bedeutung sind.

* Dies bestätigt die Vermutung, daß die Farbdias zur Gewöhnung nicht geeignet sind.

** M.E. hängen diese Wörter stärker zusammen als die Wörter der zweiten Liste.

Hier trat nicht das von WALKER postulierte Phänomen auf, wonach stark erregende Items kurzfristig schlechteres Behalten bewirken. SCHÖNPFUG & BEIKE prüften nur das Kurzzeitgedächtnis, nicht das Langzeitgedächtnis.

Bei nahezu allen anderen Free Recall Untersuchungen waren fast immer die stark erregenden Wörter den schwach erregenden Wörtern überlegen. Als Beispiel die Untersuchung von MALTZMAN, KANTOR & LANGDON (1966) zum inzidentellen Lernen an je 40 männlichen und weiblichen Studenten. Sie boten alle 10 Sekunden eines von 16 Wörtern (die 8 stark erregenden und die 8 schwach erregenden Wörter von WALKER & TARTE 1963) in einer von zwei Reihenfolgen über Kopfhörer dar. Die Gedächtnisprüfung erfolgte entweder unmittelbar oder 30 Minuten später. Von den stark erregenden Wörtern wurden unmittelbar durchschnittlich 5.82 und 30 Minuten später 4.50 erinnert, von den schwach erregenden Wörtern 3.15 und 2.62. Die stark aktivierenden Wörter wurden sowohl bei der unmittelbaren Prüfung als auch später häufiger reproduziert, was im Gegensatz zur Theorie von WALKER steht. Die Hypothese von MALTZMAN, KANTOR & LANGDON lautet: Das Auftreten von Orientierungsreaktionen erleichtert die Aufnahme und das Behalten der Wörter. Stark erregende Wörter sind leichter kategorisierbar, dies verbessert das Lernen.

KAPLAN & KAPLAN kritisieren die a priori Klassifikation der Wörter, da ein Item je nach Vp und je nach Situation eine beträchtliche Variation in den EDR aufweisen kann.*

KAPLAN, KAPLAN & SAMPSON (1968) führten auch eine Untersuchung zum Free Recall durch. 40 Studenten, die sich 24 Items einprägen sollten**, wurden unmittelbar und 30 Minuten nach

* Diese Aussage bestätigt die Annahme, daß bei ihren Versuchen eine Konfundierung der EDR mit der Position vorliegt. Die großen Unterschiede erklären sich daraus, daß die Wörter pro Vp in der Reihenfolge variierten und die Anfangsitems starke Erregung hervorriefen.

** Davon 12 als Wörter und 12 als Zeichnungen; Diawechsel alle 4 Sekunden; dazwischen Farbendias.

der Darbietung hinsichtlich ihrer Behaltensleistung geprüft. Für jede Vp wurden zwei EDR-Werte berechnet, einer für die Items, die vergessen wurden (Items, die die Vp unmittelbar erinnerte, aber nicht mehr 30 Minuten später) und eine für 'reminiscente' Items (Items, die die Vp nicht sofort, aber 30 Minuten später erinnerte). Der durchschnittliche EDR-Ausschlag für die 'reminiszenten' Items ist größer als für die vergessenen. Dieses Ergebnis spricht nach Ansicht der Autoren für die Theorie von WALKER.

SAMPSON (1969), der den Versuch* nochmals an 72 Studenten wiederholte und zusätzlich eine inzidentelle Lernbedingung miteinbezog, erhielt keine ähnlich klaren Verhältnisse mehr. "The situation with respect to arousal factors is considerably less clear." (SAMPSON, S.222)

Eine weitere Studie zum Free Recall (unabsichtliches Lernen) stammt von CORTEEN (1969). Er testet die Hypothese, daß die Größe der phasischen Leitfähigkeitsreaktion auf ein Item positiv verbunden ist mit der späteren Reproduktion des Items. Items mit größeren Reaktionen produzieren lebensfähigere dynamische Spuren, die über längere Zeit bestehen bleiben. CORTEEN verwendete drei verschiedene Wortlisten, die ersten beiden enthielten je 21 Wörter, die dritte 15 Wörter unterschiedlicher Häufigkeit. Voraus gingen fünf Übungswörter. Die Wörter wurden akustisch dargeboten. Die Vp sollte über die Bedeutung der Wörter nachdenken. Das nächste Wort wurde erst gebracht, wenn die Hautleitfähigkeit wieder auf die Basislinie zurückgekehrt war. Das führte dazu, daß ziemlich viel Zeit zwischen den einzelnen Wörtern verging, vor allem bei den Übungswörtern (10 bis 15 Minuten insgesamt), bei den Testwörtern durchschnittlich 30 bis 60 Sekunden pro Wort. Nach der Abnahme der Elektroden wurde mit Gruppe 1 (18 Vpn) die Behaltensprüfung von drei Minuten durchgeführt, mit der Gruppe 2 (23 Vpn) 20 Minuten danach und mit Gruppe 3 (19 Vpn) zwei Wochen später. Die Worthäu-

* Retentionsprüfung diesmal unmittelbar oder einen Tag später.

figkeit hatte keinen Einfluß auf die Reproduktion. Die EDR-Werte (log change in conductance) wurden pro Vp in T-Werte transformiert und über die Vpn zusammengefaßt. Die biseriellen Korrelationen zwischen den EDR-Werten und der Behaltensleistung (erinnert - nichterinnert) betrugen für die

Gruppe 1: $r_b = .13$ $\sigma_{rb} = .064$ ($p < .05$)

Gruppe 2: $r_b = .23$ $\sigma_{rb} = .056$ ($p < .01$)

Gruppe 3: $r_b = .40$ $\sigma_{rb} = .071$ ($p < .01$).

Die Ergebnisse unterstützen laut CORTEEN nicht die Theorie der Michigangruppe; die Korrelation hätte für Gruppe 1 negativ sein müssen. Die Beziehung zwischen EDR und Behalten wird mit zunehmender Zeit enger. Das langfristige Behalten ist eine Funktion der Stärke der ursprünglichen dynamischen Spur. Die Stärke der Spur hängt von der Erregung bei der Spurenbildung ab.

CORTEEN gehört zu den wenigen, die Übungswörter präsentierten. Ein entscheidender Nachteil seiner Untersuchung ist aber die unterschiedlich lange Darbietungszeit der einzelnen Items. Bei einem erregenden Wort war die Zeitspanne länger als bei einem schwach erregenden Wort. Man könnte deshalb die Zeit als Ursache für das bessere Behalten heranziehen. Es ist naheliegend, daß eine Sache, der man mehr Zeit widmet, auch besser behalten wird. Ob die Erregung oder die Zeit zum besseren Behalten der erregenden Wörter führte, läßt sich an dieser Untersuchung nicht feststellen.

Denselben Fehler, daß nämlich die Zeitdauer der Beschäftigung mit einem Item nicht kontrolliert wurde, begingen auch MANNING & JULIAN (1975). 15 weibliche und 8 männliche Studenten im Alter von 19 bis 35 Jahren sollten 30 Wörter lernen; sie erhielten je 90 Sekunden für das Lesen der Wortliste und je 120 Sekunden für die Behaltensprüfung. Drei Lernversuche wurden durchgeführt, wobei jedesmal die Reihenfolge der Wörter variierte. Je fünf Wörter gehörten einer der folgenden Kategorien an: 'sex, bathroom, violence' (erregende Wörter); 'reading material, fruit, beverages' (neutrale Wörter). Die Wörter waren bzgl. Wortlänge und Vertrautheit paralleli-

siert. Über die Versuche nahm die Anzahl der erinnerten Wörter zu. Die erregenden Wörter wurden besser erinnert als die neutralen. Eine genauere Analyse ergab, daß vor allem die Tabuwörter ('sex' und 'bathroom' Kategorie) überlegen waren. Die Gewaltitems unterschieden sich nicht von den neutralen. Das bessere Abschneiden der Tabuwörter ließ sich nicht auf ein stärkeres Clustering zurückführen. Die Untersuchung gibt einen Hinweis darauf, daß die Behaltensleistung vermutlich eher mit den EDR-Werten zusammenhängt als mit der subjektiven Beurteilung der Vpn.*

MANNING führte mit GOLDSTEIN (1976) ein weiteres Experiment (mit vier Lernversuchen) durch, bei dem die Präsentationszeit pro Item konstant war (entweder 1 sec oder 4 sec). Zusätzlich wurde noch die Zusammenstellung der Wörter variiert, entweder wurden alle Wörter einer Kategorie hintereinander (als Block) gebracht oder über die Liste verstreut. 40 Wörter (je 10 aus den Kategorien : Sexualität, Gewalt, Lesematerial und Getränke) wurden mit Hilfe einer Gedächtnistrommel 64 Studentinnen im Alter von 17 bis 40 Jahren gezeigt. Die Varianzanalyse erbrachte eine signifikante Überlegenheit der Blockliste und eine signifikante Überlegenheit der 4 sec Präsentation, ebenso ein signifikantes Anwachsen über die Versuche. Wieder schnitten die Wörter aus dem Bereich 'Sexualität' besser ab. Der Effekt war beim 1.Lernversuch und bei den nicht blockweise angeordneten Wörter am stärksten, ebenso bei der schnelleren Präsentationsrate (hier aber nicht signifikant). MANNING & GOLDSTEIN vermuten, daß Tabuwörter Items von hoher Priorität ('high priority') sind, die mehr von der vorhandenen Prozeßkapazität verbrauchen.

Abschließend sei noch die Untersuchung zum unabsichtlichen Lernen von WARREN & HARRIS (1975) angeführt, die auch eine Langzeitgedächtnisprüfung miteinbezogen und die Klassifika-

* In dieser Untersuchung wurden keine EDR-Werte erhoben; die Autorinnen greifen auf die Daten einer früheren Untersuchung (MANNING & MELCHIORI 1974) zurück.

tion der Wörter anhand der EDR- und EEG-Aufzeichnungen während des Versuchs vornahmen. Zwei Wortlisten mit je 24 einsilbigen Substantiven, die homogen waren bzgl. Konkretheit, Bildhaftigkeit, Bedeutungshaltigkeit und Häufigkeit, wurden 32 Studenten akustisch dargeboten (alle 10 sec ein Wort). Die Vpn sollten an das jeweilige Wort denken und nicht die früheren Wörter wiederholen. Die ersten 4 und die letzten 4 Wörter dienten als Übungswörter, um den seriellen Positionseffekt zu minimieren. Um einer Habituation vorzubeugen, wurde über den Versuch verteilt vor einigen Wörtern ein 800 Hz Ton gegeben.* Eine EDR wurde nur dann gewertet, wenn 1 - 2 Sekunden nach der Darbietung eines Wortes die Amplitude 0.25 mV überschritt. Ebenso wurde gewertet, ob ein Alpha-Block (AB) vorlag. Die dreiminütige Gedächtnisprüfung wurde entweder unmittelbar oder mit einer Verzögerung von 45 Minuten durchgeführt. Alpha-Blocks und elektrodermale Reaktionen kamen nicht oft gemeinsam vor, was für eine gewisse Unabhängigkeit der beiden Reaktionen spricht.** Der Anteil der physiologisch erregenden Wörter (sowohl bzgl. AB als auch EDR) ist beim kurzfristigen Behalten größer. Bei der Langzeitgedächtnisprüfung ergab sich kein signifikanter Unterschied. Gegenüber der Kurzzeitgedächtnisprüfung wurden hier sogar mehr physiologisch nicht-erregende Wörter erinnert. Dies ist eine der ersten Untersuchungen, die hinsichtlich des Langzeitgedächtnisses ein schlechteres Abschneiden der erregenden Wörter aufweist. Beim Kurzzeitgedächtnis sind vor allem die Wörter überlegen, die sowohl eine EDR als auch einen Alpha-Block hervorrufen, ebenfalls die Wörter, die nur eine EDR bewirken.

* M.E. wirkt dieser zusätzliche Reiz störend. Da Töne EDR hervorrufen können, besteht die Gefahr einer Vermischung der physiologischen Reaktionen.

** WARREN & HARRIS rechnen den Alpha-Block einem Aufmerksamkeits-Arousal-System zu, die EDR einem Affekt-Arousal-System.

Zusammenfassend läßt sich feststellen, daß die bisherigen Studien uneinheitliche Ergebnisse zeigen. Eine Reihe möglicher Fehlerquellen erlaubt keine sicheren Aussagen (Positionseffekte, unterschiedliche Darbietungszeit pro Item, Art der Itemklassifikation, Interitemassoziationen, andere das Lernen beeinflussende Wortcharakteristika).*

Um den Einfluß des emotionalen Erregungspotentials von Wörtern auf das verbale Lernen ermitteln zu können, muß gewährleistet sein, daß keine Konfundierung mit anderen Wortcharakteristika besteht, die für die Lernleistung effektiv sind.

Sieht man die Literatur zum verbalen Lernen durch, so stößt man auf verschiedene Wortcharakteristika, z.B. Bedeutungshaltigkeit, Konkretetheit, Bildhaftigkeit, Vorkommenshäufigkeit, Wortlänge, Wortart, von denen ein Einfluß auf das Lernen angenommen wird.

Im folgenden werden diese Wortparameter kurz behandelt. Das Hauptgewicht wird dabei auf Free Recall Studien gelegt, da im experimentellen Teil dieser Arbeit die Methode des Free Recall verwendet wird.

3.2. Bedeutungshaltigkeit (meaningfulness m, m', av, AF)

3.2.1. Operationalisierungen

Die Bedeutungshaltigkeit wird auf verschiedene Arten operationalisiert, das gemeinsame an den Verfahren ist, daß sie auf Assoziationen beruhen. Ein Wort, das mehr Assoziationen hervorruft, gilt als bedeutungshaltiger.

* BOCK (1978 a), der Untersuchungen zu 'Arousal und Gedächtnisleistung' referiert, kommt in seiner Zusammenfassung zu dem Ergebnis: "Stark aktivierende Items werden kurzfristig zumeist schlechter behalten als schwach aktivierende. ... Längerfristig werden sie jedoch besser behalten als schwach aktivierende Items." (S.17).

Eine solche Aussage ist m.E. wegen der angeführten Störfaktoren und der unterschiedlichen Ergebnisse nicht zulässig.

Die erste systematische Arbeit zur Bestimmung des Assoziationswertes (association value, av) von 'sinnlosen' Silben stammt von GLAZE (1928, zit. nach HALL 1971). 15 Vpn wurden ca. 2000 Silben (CVC) gezeigt (pro Silbe 2 - 3 sec). Die Vpn sollten jeweils angeben, ob die Silbe eine Assoziation bewirkt. Den Assoziationswert (av) erhält man, indem man den Prozentsatz der Vpn berechnet, bei denen eine Silbe eine Assoziation auslöste.

$$av = \frac{n_A}{n} \cdot 100$$

wobei n_A = Anzahl der Vpn, die eine Assoziation angeben
 n = Gesamtzahl der Vpn

Der Wert für av kann von 0% bis 100% variieren.

Ein ähnliches Maß ist das von ARCHER (1960, zit nach HALL). Hier bestimmt ebenfalls die Prozentzahl der Vpn (mit 'ja'-Antwort) die Bedeutungshaltigkeit von CVC-Silben. Die Silbe wird 4 Sekunden dargeboten, und die Vp soll sich fragen, 1) ob die Silbe ein Wort ist, 2) ob sie wie ein Wort klingt, 3) ob sie an ein Wort erinnert, und 4) ob man sie in einem Satz verwenden könnte. Falls eine der Fragen mit 'ja' beantwortet werden kann, soll die Vp 'ja' sagen.

Das Maß von TAYLOR & KIMBLE (1967, zit. nach HALL) gibt den Anteil der Vpn in Prozent an, die in weniger als 2.5 sec auf ein Wort (Substantiv) mit einer Assoziation reagieren.

Die eben beschriebenen Methoden sind in erster Linie für Silben anwendbar und nur bedingt zur Messung der Bedeutungshaltigkeit von Wörtern geeignet.

Für die Bestimmung der Bedeutungshaltigkeit von Wörtern stehen zwei unterschiedliche Methoden zur Verfügung, a) die Produktionsmethode und b) die Schätzmethode.

a) Produktionsmethode

Bei der Produktionsmethode werden die Vpn aufgefordert, innerhalb eines begrenzten Zeitintervalls so viele Reaktionen abzugeben, wie der Stimulus hervorruft. Den Meßwert bildet

dann die durchschnittliche Anzahl der Reaktionen innerhalb dieser Zeitspanne.

Der bekannteste Index ist der Index m (meaningfulness) von NOBLE (1952). Er wird folgendermaßen ermittelt: Die Vpn haben 60 sec Zeit (manchmal auch nur 30 sec, bei SPREEN & SCHULZ 1966; PAIVIO, YUILLE & MADIGAN 1968), in der sie ihre Assoziationen auf das Stimuluswort notieren. Das arithmetische Mittel der Anzahl der genannten Assoziationen auf ein Wort ergibt dessen m -Wert.

$$m = \frac{\sum A}{n}$$

wobei A = Assoziationen

n = Anzahl der Vpn

Der m -Wert kann im Bereich von 0 bis ca. 15 schwanken. Die Messung von m ist sehr reliabel (HALL 1971).

Ein ähnliches Maß ist die 'assoziative Häufigkeit' (associative frequency, AF). Die Vpn schreiben hier nicht die Assoziationen auf, sondern sie machen nur Striche für jede ausgelöste Assoziation. Pro Wort stehen 10 Sekunden zur Verfügung (VICORY & ASHER 1966).

$$AF = \frac{\sum \text{Striche}}{n}$$

Der Vorteil dieser Methode gegenüber m liegt im geringeren Zeitaufwand.

Das Maß N (HOWE 1972) wird ebenfalls mit Hilfe der Produktionsmethode gewonnen. Im Gegensatz zu m und AF , wo es um die Anzahl der Assoziationen pro Vp geht, handelt es sich hier um die Zahl verschiedener Assoziationen innerhalb einer Gruppe von Vpn, wobei jede Vp nur eine Assoziation äußert. Haben die Vpn dieselbe Assoziation auf ein Wort, so ist der Wert N klein, das Wort hat eine sehr homogene Bedeutung. N und m sind invers aufeinander bezogen. Nach HOWEs Ansicht ist das Maß N besser als m . Lernen wird als Funktion von $1/N$ angesehen. (Lernen = $f(1/N)$.) Je kleiner N , desto schneller erfolgt das Lernen.

b) Schätzverfahren

Da die Produktionsmethode sehr aufwendig ist, suchte man nach einem schnelleren Verfahren, das man in der Schätzmethode fand. Anhand von Ratingskalen schätzt die Vp, wieviele Assoziationen ein bestimmtes Wort bei ihr hervorrufen kann. Meist werden 5-Punkte Skalen verwendet, manchmal auch 7- (FRINCKE 1968) oder n-stufige.

Beispiel für eine 5-Punkte Skala (NOBLE 1961) mit den Markierungen: 1 keine (none = N), 2 unterdurchschnittlich (below average=BA), 3 durchschnittlich (average = A), 4 überdurchschnittlich (above average = AA), 5 sehr viele (very many = VM).

Aus den Schätzurteilen kann man drei Maße ableiten: a, a' und m'. Das Maß a entspricht dem Assoziationswert (av) von GLAZE. Man bestimmt hier die Prozentzahl der Vpn, die einen Wert zwischen 2 und 5 ankreuzen.

$$a = \frac{n - f_N}{n}$$

wobei n = Anzahl der Vpn

f_N = Häufigkeiten der Antworten in der Kategorie N.

Das Maß a' gibt den Mittelwert der Schätzurteile an.

$$a' = \frac{1(f_N) + 2(f_{BA}) + 3(f_A) + 4(f_{AA}) + 5(f_{VM})}{n}$$

Der Wert m' (scaled meaningfulness) wird berechnet durch Transformation von a' auf eine Standardwertskala mit der Methode der sukzessiven Intervalle. (Formel siehe NOBLE 1961, S.508)

Eine andere Ratingskala ist die Skala AD (associative difficulty) von BROWN & URE (1969). Auf einer 7-Punkte Skala mit den Polen 1 (= leicht zu assoziieren) und 7 (= schwierig zu assoziieren) sollen die Vpn einschätzen, wie schwer es ihnen fällt, auf ein Wort zu assoziieren. BROWN & URE geben zwei Reliabilitätskoeffizienten an: .969 (Wiederholungsmessung)

und .909 (Übereinstimmung zweier Beurteilerstichproben).

Auch die anderen Maße der Bedeutungshaltigkeit sind nach bisheriger Erfahrung sehr reliabel. Alle Maße weisen hohe Interkorrelationen auf (.70 bis .95, HALL 1971 und BROWN & URE 1969).

3.2.2. Einfluß auf das verbale Lernen

Eine der frühesten Studien zum Einfluß der Bedeutungshaltigkeit auf das Lernen stammt von McGEOCH (1930, zit. nach McGEOCH & IRION 1952). Er benutzte 10 Items umfassende Listen, die entweder aus dreibuchstabigen Wörtern oder Silben unterschiedlichen assoziativen Wertes (av 0%, 53%, 100%) zusammengesetzt waren. Die Vpn hatten je zwei Minuten Zeit, um die Liste zu studieren und zwei Minuten zur anschließenden Reproduktion. Die Anzahl der erinnerten Items variierte direkt mit dem assoziativen Wert. (Tabelle 9)

Material	Mittelwert	G
Wörter	9.11	1.12
100% Silben	7.53	1.96
53% Silben	6.41	2.37
0% Silben	5.09	2.60

Tabelle 9: Reproduktionsmenge unterschiedlich bedeutungshaltigen Materials. (McGEOCH & IRION, S.470)

Die Behaltensleistung nahm mit der Größe des Assoziationswertes zu. Am besten wurden die Wörter erinnert. Es zeigte sich eine positive Beziehung zwischen der Bedeutungshaltigkeit und der Menge des gelernten Materials. McGEOCH & IRION erklären den erleichternden Effekt der Bedeutungshaltigkeit damit, daß die Vp schon vorher mehr über das bedeutungshaltige Material weiß und dadurch Vorteile aus Transfereffekten bezieht.

Bei der Prüfung, welche Rolle die unterschiedliche Bedeutungshaltigkeit von Wörtern beim Lernen spielt, wurden meist nur zwei Ausprägungen von m verglichen, Wörter mit hohem m-Wert

(high m = H) und Wörter mit niedrigem m-Wert (low m = L). Dabei fand sich eine Überlegenheit der Wörter mit hohem m-Wert (SEAMAN & MURRAY 1976; PAIVIO, YUILLE & ROGERS 1969).

In der Untersuchung von LEY & KARKER (1977) korrelierte die skalierte Bedeutungshaltigkeit (m') signifikant mit der Recallhäufigkeit.

Der das Lernen fördernde Effekt wurde öfters repliziert und trat auch beim seriellen Lernen und beim Paarassoziationslernen auf. Bei letzterem wird das Lernen vor allem verbessert, wenn das Response-Item einen hohen m-Wert aufweist, aber auch in der Stimulus-Position erleichtert die Bedeutungshaltigkeit das Lernen. CIEUTAT, STOCKWELL & NOBLE (1958, zit. nach KINTSCH 1977) erhielten folgende Reihenfolge bzgl. der Menge richtiger Antworten: H-H, L-H, H-L, L-L. Die meisten richtigen Antworten bekamen sie, wenn sowohl das Stimulus- als auch das Response-Wort einen hohen m-Wert hatten, die schlechteste Leistung ergab sich, wenn beide Wörter einen niedrigen m-Wert hatten.

Die Bedeutungshaltigkeit scheint das verbale Lernen zu erleichtern. Dieser Faktor spielt besonders beim Lernen sinnarmer Silben eine Rolle.

3.3. Konkretheit - Abstraktheit (concreteness C, specificity S)

3.3.1. Operationalisierungen

Es lassen sich zwei Arten von Konkretheit - Abstraktheit unterscheiden:

- 1) Die Bezogenheit auf sinnliche Erfahrung. Was mit den Sinnen wahrgenommen werden kann, ist konkret (z.B. 'Tisch'); was nicht mit den Sinnen wahrgenommen werden kann, ist abstrakt (z.B. 'Unabhängigkeit').
- 2) Die Unterscheidung in spezifische Begriffe und in allgemeine Begriffe (Gattungsnamen) - (specific vs. generic). Oberbegriffe (z.B. 'Möbel') sind abstrakter als spezifische Bezeichnungen (z.B. 'Stuhl').

Die Konkretheit oder Abstraktheit eines Wortes wird durch Einschätzungen erfaßt. Entweder wird nur eine Einteilung in die Kategorien 'abstrakt' und 'konkret' vorgenommen (DUKES & BASTIAN 1966) oder der Grad der Konkretheit (oder Spezifität) wird auf einer n-Punkte Skala (meist 7 Punkte, manchmal auch 5 Punkte) eingestuft (SPREEN & SCHULZ 1966; FRINCKE 1968; PAIVIO, YUILLE & MADIGAN 1968; BROWN & URE 1969).

GORMAN (1961, zit. nach HALL 1971) ließ 1791 Substantive von zwei Beurteilern in die beiden Kategorien 'konkret' und 'abstrakt' einteilen. Die Übereinstimmung zwischen den beiden Beurteilern war sehr hoch, bei 96% der Wörter wurde die gleiche Kategorie gewählt.

Auch die Einschätzungen auf den mehrstufigen Ratingskalen sind sehr zuverlässig; sie weisen hohe Reliabilitätskoeffizienten auf (.94 bis .984, BROWN & URE 1969; PAIVIO, YUILLE & MADIGAN 1968).

Die Skalen C (concreteness) und S (specificity) korrelieren nicht so hoch miteinander ($r = .626$, SPREEN & SCHULZ 1966, 329 Wörter; $r = .73$, BROWN & URE 1969, 36 Wörter). Die beiden Skalen erfassen also etwas unterschiedliche Aspekte. PAIVIO (1967, zit. nach HALL 1971, S.65) demonstrierte, daß Wörter beide Attribute (C und S) in unterschiedlicher Ausprägung aufweisen können.

Beispiel: Werkzeug: konkret - allgemein
 Hammer : konkret - spezifisch
 Emotion: abstrakt - allgemein
 Kummer: abstrakt - spezifisch

3.3.2. Einfluß auf das verbale Lernen

Es herrscht die allgemeine Überzeugung, daß konkrete Wörter leichter gelernt werden als abstrakte. Schon BUSEMANN (zit. nach FRÖBES 1923, S.536) stellte fest, daß konkrete Substantive leichter zu lernen sind als abstrakte.

Die Untersuchung von STOKE (1929, zit. nach McGEACH & IRION

1952) erbrachte, daß am besten konkrete Wörter, dann onomatopoeische Wörter und am schlechtesten abstrakte Wörter behalten werden.

In der Untersuchung von DUKES & BASTIAN (1966) erhielten die Vpn Lernlisten, die 10 abstrakte und 10 konkrete Substantive enthielten, die bzgl. der Bedeutungshaltigkeit und Vorkommenshäufigkeit parallelisiert waren. Es wurden signifikant ($p < .05$) mehr konkrete Wörter erinnert. Der Effekt trat bei den Wörtern mit hoher Vorkommenshäufigkeit auf.

GORMAN (1961, zit. nach DUKES & BASTIAN) zeigte, daß konkrete Wörter eher wiedererkannt werden als abstrakte Wörter.

Auch beim Paarassoziationslernen erleichtert die Konkretheit eines Wortes das Lernen, vor allem wenn das Stimuluswort konkret ist (PAIVIO 1965). Am leichtesten zu lernen ist die Kombination C-C, C-A, dann A-C und am schwierigsten A-A (A = abstrakt, C = konkret).

3.4. Bildhaftigkeit (imagery I, vividness V)

3.4.1. Operationalisierungen

Unter Bildhaftigkeit (imagery) versteht man die Leichtigkeit oder Deutlichkeit, mit der ein bestimmtes Wort ein Bild oder eine visuelle Vorstellung hervorruft. Anstelle von 'imagery' taucht auch die Bezeichnung 'vividness' (Lebendigkeit) auf.

Die Bildhaftigkeit wird auf einer mehrstufigen Ratingskala eingeschätzt, z.B. auf einer 5-Punkte Skala (PAIVIO 1965, 1968: 'very easy - very difficult image scale') oder auf einer 7-Punkte Skala (mit den Polen 'clear image - unclear image', FRINCKE 1968; 'low imagery - high imagery, PAIVIO, YUILLE & MADIGAN 1968).

Wörter, die schnell und leicht eine visuelle Vorstellung auslösen oder ein deutliches, klares Bild liefern, bekommen einen hohen I-Wert.

Die Bildhaftigkeitsskalen weisen ähnlich hohe Reliabilitätskoeffizienten auf wie die Konkretheitsskalen. Zwischen der Bildhaftigkeit (I) und der Konkretheit (C) ergeben sich relativ hohe Korrelationen. Bei PAIVIO, YUILLE & MADIGAN (1968) beträgt der Korrelationskoeffizient .83, bei FRINCKE (1968) .92. Der enge Zusammenhang zwischen I und C ist verständlich, konkrete Wörter rufen eher eine visuelle Vorstellung hervor als abstrakte Wörter. PAIVIO (1969, S.245) schreibt: "Except for certain unusual items ..., however, the two scales appear largely to be alternative measures of the same variable." Die Ausnahmen, Wörter also, bei denen die Einschätzungen bzgl. der Konkretheit und der Bildhaftigkeit divergieren, betreffen folgende Fälle: Wenn ein Wort um mehr als zwei Skaleneinheiten (auf 7-Punkte Skalen) konkreter beurteilt wird als bildhaft, handelt es sich um ein sehr ungewöhnliches, seltenes Wort. Wird ein Wort aber um mehr als zwei Skaleneinheiten bildhafter als konkret eingeschätzt, bedeutet das, daß dieses Wort mit hoher Wahrscheinlichkeit eine stark emotionale oder evaluative Bedeutung hat. (PAIVIO, YUILLE & MADIGAN 1968)

3.4.2. Einfluß auf das verbale Lernen

Die Bildhaftigkeit ist hoch korreliert mit der Leichtigkeit des Lernens (PAIVIO 1971, zit. nach KINTSCH 1977).

In der Free Recall Untersuchung von TULVING, McNULTY & OZIER (1965, zit. nach HALL 1971) zeigte sich, daß am leichtesten das Material gelernt wird, das die höchsten V (vividness)-Werte aufwies. Die Vpn sollten in acht Versuchsdurchgängen drei Listen à 16 Wörter lernen, die sich in den V-Werten unterschieden (hoch, mittel, niedrig), bzgl. m aber ausgeglichen waren.

In der Untersuchung von PAIVIO, YUILLE & ROGERS (1969) wurde ebenfalls die Überlegenheit von 'high-I' Wörtern gegenüber 'low-I' Wörtern beim Free Recall und beim seriellen Lernen demonstriert. Beim Free Recall übertraf der Effekt

von I (m konstant) signifikant den Effekt von m (I konstant). Die Bildhaftigkeit wird, was das Lernen betrifft, als wichtigere Variable angesehen als die Bedeutungshaltigkeit.

Auch in Paarassoziationsstudien (PAIVIO & YUILLE 1967, SMYTHE & PAIVIO 1968) erwies sich, daß I wirksamer ist als m. Der Effekt von I ist wie bei der Konkretetheit auf der Stimulusseite größer. Die zuletzt erwähnte Studie erbrachte sogar eine negative Beziehung für m bei Items mit niedrigen I-Werten. Hier wurden die Items mit niedrigen m-Werten besser gelernt.

In der Free Recall Untersuchung von HOHN, JOHNSEN & TRACY (1976) wurden auch die 'high-I' Wörter öfter erinnert als die 'low-I' Wörter sowohl bei der unmittelbaren Behaltensprüfung als auch eine Woche später; der Unterschied war aber nicht statistisch signifikant.

3.5. Vorkommenshäufigkeit, Vertrautheit (frequency F, familiarity f)

3.5.1. Operationalisierungen

Zwei Arten der Bestimmung lassen sich unterscheiden:

- a) die objektive Häufigkeit, die auf Wortzählungen basiert,
- b) die subjektive Häufigkeit, die auf Einschätzungen beruht.

a) objektive Häufigkeit (frequency F)

Die Vorkommenshäufigkeit eines Wortes wird anhand von umfangreichen Wortzählungen bestimmt. Es werden die Häufigkeiten ausgezählt, mit der die einzelnen Wörter in schriftlichem oder gesprochenem Material auftreten.

Die meisten Zählungen stützen sich auf die geschriebene Sprache. Im englischen Sprachraum ist die THORNDIKE-LORGE Häufigkeitszählung von 1944 am bekanntesten, ihre Werte werden am häufigsten verwendet. Es wurden Texte ausgewertet, die insgesamt 5 Millionen Wörter umfaßten. Die amerikanische Worthäufigkeitszählung von KUCERA & FRANCIS (1967, zit. nach

HALL 1971) ist jüngeren Datums; hier wurden nur Texte nach 1961 ausgezählt, die Stichprobe enthielt etwas über 1 Million Wörter.

Von HOWES (1966, zit. nach HALL) gibt es Häufigkeitswerte für die gesprochene englische Sprache. 1/4 Million Wörter aus 50 Interviews zwischen 1960 und 1965 wurden dabei verarbeitet.

Für den deutschen Sprachraum sind die Worthäufigkeitszählungen von KAEDING (1898) und von MEIER (1967) zu erwähnen. Bei KAEDING wurden fast 11 Millionen deutsche Wörter ausgezählt. Erfasst wurde der geschriebene Wortschatz des gebildeten Erwachsenen gegen Ende des letzten Jahrhunderts. Die 'Deutsche Sprachstatistik' von MEIER basiert zum größten Teil auf dem Häufigkeitswörterbuch von KAEDING.

Zur gesprochenen Sprache gibt es zwei Zählungen, das 'Rangwörterbuch hochdeutscher Umgangssprache' von H.H.WÄNGLER (1963) und von J.A. PFEFFER (1964) 'Grunddeutsch' (600 000 Wörter aus Interviews und Erzählungen).

Die Problematik der Häufigkeitstabellen liegt darin, daß die Aktualität der Werte begrenzt ist. Die Sprache ist dauernden Wandlungen unterworfen, deshalb stimmen die angegebenen Zahlenverhältnisse häufig nicht mehr. Bei manchen Wörtern zeigt sich ein schneller Wandel, z.B. waren die Wörter 'nuklear', 'Stereo', 'Diskothek' vor dem 2. Weltkrieg ziemlich selten (TRYK 1968, zit. nach HALL, S.55), heute dagegen werden sie häufiger verwendet. Andere Wörter wiederum geraten weitgehend in Vergessenheit; sie sind kaum noch aktuell, z.B. 'Kaiser', 'Krone' (MARX 1978, S.438). Die ständigen Veränderungen in Gesellschaft und Umwelt lassen immer wieder neue Wörter entstehen, die in den älteren Zählungen daher noch nicht auftauchen können (z.B. 'Atomkraftwerk', 'Computer', 'Fernsehen').

Da die meisten Zählungen nur schriftliches Material analysierten und die gesprochene Sprache nicht berücksichtigten, verzerrt sich das Bild. Ebenso kann die Textauswahl in be-

stimmten Punkten sehr einseitig sein, so daß manche Wörter unterrepräsentiert sind. Sicher ist, daß Wörter aus den sogenannten 'Tabubereichen' oft nicht vertreten sind, obwohl es sich hier kaum um unbekannte Wörter handelt. Natürlich trägt zur Bekanntheit eines Wortes auch die Häufigkeit bei, mit der an ein Wort gedacht wird. Dies geht in die Zählungen nicht mit ein.

Die vorliegenden Sprachzählungen spiegeln nur Gruppenwerte wider, deshalb lassen sich an ihnen die spezifischen Werte einer bestimmten Person nicht ablesen.

b) subjektive Häufigkeit oder Vertrautheit (familiarity f)

Zur Feststellung des Bekanntheitsgrades von Wörtern bedient man sich nicht nur der objektiven Wortzählungen, sondern erhebt auch Urteile der Vpn über die Vertrautheit von Wörtern. Die Vpn schätzen auf Ratingskalen ein, wie bekannt ihnen ein Wort ist oder wie häufig sie ein Wort verwenden. Verschiedene Ratingskalen werden benutzt:

5-stufige (NOBLE 1953*, MANNING & JULIAN 1975*)

7-stufige (HAAGEN 1949**, PAIVIO 1968**, FRINCKE 1968*)

9-stufige (PAIVIO 1965**).

3.5.2. Einfluß auf das verbale Lernen

Die Häufigkeit eines Wortes kann zum Lernerfolg beitragen. In der Literatur ist vom 'Häufigkeitsparadoxon' die Rede (GREGG 1976). Dies bedeutet, daß beim Free Recall die bekannten Wörter (= Wörter hoher Häufigkeit) besser behalten werden als die seltenen Wörter, während bei der Wiedererkennungsmethode der umgekehrte Fall eintritt. RUNDLE (1970) erhielt in ihren Untersuchungen zum Free Recall und zur Rekognition diese Ergebnisse.

HALL (1954, zit. nach HALL 1971) fand in seiner Free Recall

* Häufigkeit des Umgangs

** Bekanntheit oder Vertrautheit

NOBLE 1953 und HAAGEN 1949, zit. nach HALL 1971.

Untersuchung eine monotone Beziehung zwischen dem Häufigkeitswert eines Wortes (nach THORNDIKE-LORGE T-L) und der Anzahl der erinnerten Wörter. Vier Wortlisten à 20 Wörter, die unterschiedliche Häufigkeitswerte aufwiesen, wurden fünfmal dargeboten und ergaben folgende Behaltensleistung:

<u>T-L Häufigkeit</u>	<u>Durchschnittliche Reproduktionsmenge</u>
1 pro Million	12.04
10 pro Million	13.31
30 pro Million	15.02
50 - 100 pro Million	15.04

Tabelle 10: Durchschnittliche Reproduktionsmenge bei Wörtern unterschiedlicher Häufigkeit. (HALL 1971, S.98)

Die Behaltensleistung für seltenere Wörter ist schlechter als für geläufigere. Von einem bestimmten Häufigkeitswert (30 pro Million) an zeigen sich keine Unterschiede mehr.

Ähnliche Ergebnisse erhielten auch andere Autoren (z.B. SUMBY 1963, zit. nach LLOYD 1964).

Der Häufigkeitseffekt konnte aber nicht immer nachgewiesen werden. PAIVIO (1968) und FRINCKE (1968) fanden keinen Zusammenhang zwischen T-L Häufigkeit und Behalten.

HALL (1971) sieht den Grund im Fehlen deutlicher Unterschiede zwischen den Wörtern und führt die Äußerung von UNDERWOOD & SCHULZ (1960) an, die vermuten, daß auch die seltenen Wörter (in der THORNDIKE-LORGE Zählung) von den Vpn genügend oft erfahren wurden.

GREGG (1976), der einen Überblick über die Untersuchungen und theoretischen Erklärungsversuche zum Häufigkeitseinfluß bringt, stellt fest, daß sich nur in den Free Recall Untersuchungen, die homogene Listen (nur Wörter einer Häufigkeit) verwendeten, die Überlegenheit der Wörter hoher Häufigkeit über Wörter niedriger Häufigkeit zeigte. Die Untersuchungen, die gemischte Listen benutzten, erbrachten meist keine Unterschiede, manchmal sogar eine Überlegenheit der niedrig frequenten Wörter. Bei den Wiedererkennungsforschungen wurden sowohl bei homogenen Listen als auch bei gemischten

Listen die Wörter niedriger Auftretenswahrscheinlichkeit besser wiedererkannt als die Wörter hoher Häufigkeit. GREGG erwähnt aber auch, daß es Ausnahmen gebe.

Beim Paarassoziationslernen bedeutet eine Zunahme der Häufigkeit auf der Responseseite eine Verbesserung beim Lernen, auf der Stimulusseite ist zunächst ein Anwachsen zu beobachten, dann aber ein Abfallen. Wenn das Stimuluswort eine mittlere Häufigkeit aufweist, ist die Leistung besser als bei einem seltenen oder hochfrequenten Stimulus. (POSTMAN 1962, zit. nach HALL 1971)

In der Untersuchung von PAIVIO & MADIGAN (1970, zit. nach GREGG 1976) zeigte sich eine Interaktion zwischen Konkretheit und Häufigkeit. Bei den konkreten Wörtern wurden die Wörter hoher Häufigkeit besser reproduziert, bei den abstrakten die Wörter niedriger Häufigkeit.

3.6. Wortart

Zum Einfluß der Wortart auf das Lernen liegen bisher nur vereinzelt Untersuchungen vor. BUSEMANN (zit. nach FRÖBES 1923, S.536) vermutete, daß Adjektive und Verben schlechter gelernt werden als Substantive. HALL (1971) weist auf die Studie von EPSTEIN, ROCK & ZUCKERMAN (1960) hin, in der konkrete Substantive leichter gelernt wurden als Verben, und diese leichter als abstrakte Substantive.

Nimmt man an, daß Substantive bedeutungshaltiger sind als Verben und Adjektive, so könnte man erwarten, daß Substantive besser gelernt werden. Der Effekt wäre dann der Bedeutungshaltigkeit zuzuschreiben. MARX (1978) konnte in seiner Untersuchung zeigen, daß Substantive durchschnittlich mehr assoziative Bedeutung tragen als Adjektive und Verben vergleichbarer Auftretenshäufigkeit und vergleichbaren Informationsgehalts. Die Liste der Substantive unterschied sich in den m-Werten auf dem 1% Niveau signifikant (Vorzeichen-test) von den Listen der Adjektive und Verben, die keine signifikante Differenz aufwiesen.

Die Untersuchung von SCHUTZ & KEISLAR (1972) an kleinen Kindern ergab, daß Substantive den Verben überlegen waren und daß die Behaltensleistung bei Funktionswörtern am schlechtesten ausfiel. Die Kinder bekamen Wörter einer Klasse hintereinander vorgelesen (von zwei Items an jeweils um eines gesteigert) und sollten sie nachsprechen. Der Behaltenswert für eine Wortklasse resultierte aus der längsten Liste der Wörter dieser Klasse, die ohne Fehler wiederholt wurden. Besonders deutlich waren die Ergebnisse bei Kindern aus niedrigem sozioökonomischem Milieu.

Die Fähigkeit von Kindern, Wörter wiederzuerkennen, untersuchte KUSA (1976). Die 36 Wörter umfassende Lernliste bestand zu je einem Drittel aus Substantiven, Adjektiven und Verben. Die Rekognitionsleistung der Kinder fiel von Substantiven über Adjektive zu Verben ab.

In der Untersuchung zum Paarassoziationslernen (Wort - CVC) von GLANZER (1962, zit. nach HALL 1971) schnitten am besten die Substantive ab, am zweitbesten die Adjektive. Handelte es sich bei den Stimuli um Adverbien, Verben, Pronomen, Präpositionen oder Konjunktionen, wurden weniger richtige Paare wiedergegeben. Sah die Lernaufgabe aber folgendermaßen aus: CVC - Wort - CVC, waren die Funktionswörter den Inhaltswörtern überlegen. COFER (zit. nach HALL) führt dies auf die leichtere Aussprechbarkeit zurück.

3.7. Wortlänge

Die Wortlänge läßt sich entweder über die Anzahl der Buchstaben oder die Anzahl der Silben bestimmen. Kürzere Wörter werden vermutlich schneller gelernt als längere Wörter. Die Wortlänge hängt mit der Häufigkeit zusammen: häufige Wörter sind meist kürzer, seltene Wörter sind oft länger (ZIPF 1965).

SPREEN & SCHULZ (1966) analysierten 329 Substantive. Dabei fanden sie heraus, daß die längeren Wörter geringere Bedeutungshaltigkeit (niedriger m-Wert) und geringere Konkretheit

(und Spezifität) aufwiesen.

Die Free Recall Untersuchung von FRINCKE (1968) erbrachte eine negative Beziehung zwischen Wortlänge und Behaltensleistung ($r = -.37$, $p < .05$). Wurden die anderen Wortparameter aber konstantgehalten, war die Korrelation nicht mehr signifikant.

Manche Parameter scheinen nur deshalb einen Einfluß zu haben, weil sie mit anderen korreliert sind. Im folgenden sollen die Interkorrelationen der verschiedenen Variablen betrachtet werden, um festzustellen, welcher Faktor sich als wesentlich für das Lernen herauskristallisiert.

3.8. Interkorrelationen der verschiedenen Wortparameter

In den frühen Untersuchungen wurde meist nur ein Wortcharakteristikum herangezogen und seine Beziehung zum Lernen geprüft. Solange dabei aber nicht die anderen Wortcharakteristika kontrolliert wurden, läßt sich mit den Ergebnissen nicht viel anfangen. Es könnte sein, daß eine andere Variable, die mit der untersuchten Variablen konfundiert war, die Lernunterschiede verursachte. Um festzustellen, wie die einzelnen Wortcharakteristika zusammenhängen, sind Interkorrelationsberechnungen notwendig.

In Tabelle 11 sind Korrelationen zwischen den einzelnen Wortparameter aus verschiedenen Studien zusammengestellt:

- 32 Wörter: PAIVIO (1965, S.35)
- 96 Wörter: PAIVIO (1968, S.45)
- 925 Wörter: PAIVIO, YUILLE & MADIGAN (1968, S.7)
- 329 Wörter: SPREEN & SCHULZ (1966, S.461)
- 101 Wörter: SILVERSTEIN & DIENSTBIER (1968 S. 86)
- 650 Wörter: BROWN & URE (1969, S.234)
- 96* Wörter: NOBLE (1958)
- 48 Wörter: UNDERWOOD, EKSTRAND & KEPPEL 1965, zit. nach HALL
- 200 Wörter: HALL (1967, zit. nach HALL 1971, S.72 ff.) 1971)
- 96⁺ Wörter: NOBLE (1953/1963, zit. nach HALL 1971, S.72 ff.)
- 61 Wörter: UNDERWOOD & SCHULZ (1960, zit. nach COFER 1971, S.900)

	m'	C	S	I	V	F	f	WL	G	P	E	e
m	.92 (96) ⁺ -.03 (96)	.704 (329) .56 (925) .54 (96)	.564 (329)	.90 (32) .72 (925) .69 (96)	.61 (96)	.75 (32) .33 (925) .21 (96)	.92 (96) ⁺ .23 (96)	-.339 (329)	.11 (96)		-.26 (96)	.57 (96) * -.30 (61)
m'		-.27 (96)		-.09 (96)	.03 (96)	-.10 (96)	-.05 (96) .88 (200)		.14 (96)		.37 (96)	
AD		-.278 (650)				-.472 (650)			-.510 (650)	-.569 (650)	-.252 (650)	
C			.626 (329)	.83 (925) .78 (96)	.65 (96)	.12 (925) .01 (96) -.02 (650)	.01 (96)	-.351 (329)	.08 (96) .183 (650)	.185 (650)	-.54 (96) -.379 (650)	
S								-.300 (329)				
I					.79 (96)	.02 (96) .23 (925) .76 (32)	.09 (96)		.12 (96)		-.32 (96)	
F							.60 (96) .85 (48)		.13 (96) .239 (650)	.207 (101) .258 (650) .233 (101)	-.26 (96) .113 (650)	
f									.38 (96)		-.25 (96)	
G										.958 (650)	-.070 (650)	
P											-.058 (650)	

Tabelle 11: Interkorrelationen verschiedener Wortparameter aus unterschiedlichen Untersuchungen.

Hinter den Werten steht jeweils in Klammer die Größe der untersuchten Wortstichprobe.

(Bei m' von PAIVIO 1968 liegt eine andere Skala zugrunde mit den Polen 'meaningfulness' und 'meaningless'.)

Die Korrelationskoeffizienten unterscheiden sich je nach Art und Umfang der Wortstichprobe, die meisten fallen aber relativ hoch aus. An den Korrelationen zwischen Häufigkeit und Bedeutungshaltigkeit und zwischen Häufigkeit und Bildhaftigkeit sieht man, daß die Werte je nach Stichprobe stark variieren.*

Kein Zusammenhang zeigt sich zwischen Häufigkeit und Konkretheit und zwischen Bildhaftigkeit und der angenehmen Qualität von Wörtern.

Eine Verbindung zwischen der positiven Qualität von Wörtern (G oder P) und der Bedeutungshaltigkeit wird angenommen (JENKINS & RUSSEL 1959, zit. nach AMSTER 1964). Auf angenehme Wörter läßt sich leichter assoziieren (BROWN & URE 1969).

In der Literatur wird auch eine Beziehung zwischen der positiven Qualität von Wörtern und deren Vorkommenshäufigkeit und Vertrautheit angegeben (JOHNSON, THOMSON & FRINCKE 1960, zit. nach PAIVIO 1968). BROWN (1971) erhielt in seiner Analyse von 10 Wortlisten bei 9 eine positive Korrelation zwischen P und F.

Auffällig ist die negative Beziehung zwischen der Intensität der affektiven Tönung (E) und der Konkretheit. (Siehe dazu auch die Analyse von BROWN 1971; S.73 dieser Arbeit.) Dies steht im Gegensatz zu der Aussage von SKINNER (1957, S.158; zit. nach PAIVIO 1968, S.44), daß konkrete Begriffe eher mit spezifisch affektiven Stimuli assoziiert werden und deshalb größere emotionale Effekte haben als abstrakte Begriffe. Träfe dies zu, könnte der Effekt der Konkretheit beim verbalen Lernen in Begriffen der Emotionalität interpretiert wer-

* Wie die Verhältnisse für die Sprache an sich sind, läßt sich an diesen Werten nicht ablesen, da die Wortlisten keine repräsentative Stichprobe darstellen. Einen Hinweis liefern eventuell die Werte der umfangreichsten Stichprobe von PAIVIO, YUILLE & MADIGAN (1968), obwohl hier die Wortauswahl nicht ganz zufällig geschah. Aus der THORNDIKE-LORGE Zählung wurde nur ein Teil der Wörter per Zufall gewählt, der Rest wurde mit abstrakten Wörtern aufgefüllt, um etwa gleich viele abstrakte wie konkrete Wörter in der Stichprobe zu haben. Bei THORNDIKE-LORGE sind konkrete Wörter viel zahlreicher als abstrakte.

den (PAIVIO 1968). Versuche der Verfasserin erbrachten einen Hinweis darauf, daß konkrete Wörter eher elektrodermale Reaktionen auslösen als abstrakte. Bei SMITH & HARLESTON (1966) aber wiesen die abstrakten Wörter signifikant größere EDR auf. Mögliche Erklärungen für diesen Befund wären, a) daß abstrakte Wörter stärker affektiv getönt sind oder b) daß die EDR die größeren Schwierigkeiten beim Assoziieren widerspiegeln. Möglicherweise wären die physiologischen Reaktionen hier anders verlaufen, wenn die Vpn statt auf die Wörter zu assoziieren diese lediglich gelesen hätten.

Um der Gefahr unkontrollierter Einflüsse zu begegnen, wurde bei einer Reihe von Untersuchungen zum verbalen Lernen jeweils ein Faktor variiert, während ein anderer oder mehrere andere konstant gehalten wurden.

Bei DUKES & BASTIAN (1966), die C und F variierten* und m konstant hielten, ergab sich, daß bei Wörtern niedriger Häufigkeit kein Unterschied in der Leichtigkeit des Erinnerens zwischen A und C bestand, bei Wörtern hoher Häufigkeit mehr konkrete Wörter erinnert wurden. Bei den abstrakten Wörtern wurden mehr Wörter niedriger Häufigkeit erinnert. (Siehe Tab.12)

	abstrakt	konkret	
hohe Häufigkeit	1.67	2.37	4.04
niedrige Häufigkeit	2.37	2.41	4.78
	4.04	4.78	

Tabelle 12: Free Recall Leistung bei unterschiedlicher Konkretheit und Häufigkeit. (DUKES & BASTIAN, S.456)

Eine Interaktion zwischen C und F erhielten PAIVIO & MADIGAN (1970, zit. nach GREGG 1976). Bei den konkreten Wörtern wurden mehr Wörter mit hohem Häufigkeitswert erinnert, bei den abstrakten mehr mit niedriger Häufigkeit.

PAIVIO, YUILLE & ROGERS (1969) variierten entweder m oder I

* Die Lernliste umfaßte 10 abstrakte und 10 konkrete Wörter, die je zur Hälfte hohe bzw. niedrige Häufigkeit aufwiesen.

unter Konstanthaltung des anderen. Es zeigte sich, daß I die effektivere Variable beim Free Recall ist.

PAIVIO (1968) und FRINCKE (1968) unternahmen den Versuch, mit Hilfe der Faktorenanalyse herauszufinden, welche Maße jeweils den gleichen Aspekt erfassen und welcher Faktor den entscheidenden Einfluß auf das Lernen hat.

PAIVIO (1968) ließ 96 Substantive auf mehreren Skalen einstufen und lernen. Insgesamt wurden für die 96 Substantive 30 Variablen erfaßt, ihre Interkorrelationen (teilweise in Tab. 11 enthalten) wurden berechnet und anschließend faktorenanalysiert. Unter den Variablen befanden sich neben Paarassoziationslernenden und Free Recall-Daten (wie oft ein Wort erinnert wurde) die Werte m, m', C, S, I, V, T-L F, f, G, E, einige Skalenwerte des Semantischen Differentials und die Polaritätswerte dieser SD-Skalen**. Die Korrelationen der Variablen mit den Free Recall-Werten fallen nicht sehr hoch aus; signifikant ($p < .05$) sind nur folgende Korrelationen:

Vividness/Free Recall	r = .25
Imagery/Free Recall	r = .32
Concreteness/Free Recall	r = .23
Hardness/Free Recall	r = .25
Polarity/Free Recall	r = .28

Tabelle 12: Signifikante Korrelationen mit Free Recall.

Die Faktorenanalyse erbrachte sechs Faktoren:

- Faktor 1: Konkretheit-Bildhaftigkeit (Concreteness-Imagery);
hier lädt auch m (.64) und Lernen PA-S (.41) FR (.20)
- Faktor 2: Eindruck (Impressiveness), Interesse, Aktivität,
Emotionalität; hier lädt auch m' (.60)***
- Faktor 3: Vertrautheit (Familiarity); hier lädt auch G (.64)

* Hier handelt es sich um die beurteilte Bedeutungshaltigkeit auf einer Skala mit den Polen 'meaningful - meaningless'.

** Enthalten waren die Skalen G und E.

*** PAIVIO weist auf die Korrespondenz mit dem Activity Faktor bei OSGOOD hin. Die anderen zwei von Osgood postulierten Faktoren konnten nicht herausgearbeitet werden.

Faktor 4: Spezifität (Specificity, Preciseness)

Faktor 5: Assoziative Varietät (Associative Variety)

Faktor 6: Lernen (Learning), sowohl Paarassoziationslernen PA als auch Free Recall FR; leichte Ladung von I (.28) V (.32) und m (.30).

Faktor 2, 3, 4 und 5 weisen keine Beziehung zum Lernen auf. Faktor 1 hat eine Beziehung zum Paarassoziationslernen, wenn sich das Wort in der Stimulusposition befindet, und eine leichte Beziehung zum Free Recall. PAIVIO betont die Dominanz des 'Konkretheit-Bildhaftigkeit' Faktors für das Lernen, vor allem für das Paarassoziationslernen.

FRINCKE (1968) erhob von 70 Substantiven neben den Free Recall-Werten (Anzahl der Vpn, die das Wort erinnerten) folgende Charakteristika: m, m', C, I, T-L F, f, G (Wert aus der P und G-Skala), e-p (evaluative Polarity) und WL (Wortlänge, Anzahl der Buchstaben). Er vermutet, daß vor allem ein oder zwei Wortparameter für das Lernen grundlegend sind und die anderen nur insofern wichtig sind, als sie mit diesen zusammenhängen. Von den einzelnen Variablen korrelieren signifikant ($p < .01$) mit den Free Recall-Werten:

Imagery/FR	$r = .51$
Concreteness/FR	$r = .46$
Meaningfulness m/FR	$r = .33$
Familiarity/FR	$r = -.33$
Word length/FR	$r = -.37$

Tabelle 13: Signifikante Korrelationen mit Free Recall (wenn keine der anderen Variablen konstant gehalten wird). (FRINCKE, S.369)

Werden alle anderen Variablen konstant gehalten, korrelieren nur noch I und f signifikant ($p < .05$) mit Free Recall:

Imagery/FR	$r = .28$
Familiarity/FR	$r = -.26$

Die Faktorenanalyse* erbrachte zwei Faktoren:

* Da I und C hoch korrelierten ($r = .92$) wurde C nicht in die Analyse miteinbezogen.

Faktor 1: Bedeutungshaltigkeit-Vertrautheit (meaningful-familiarity); hier lädt auch G (.51) und e-p (.51)

Faktor 2: Bildhaftigkeit-Konkretheit (imagery-concreteness); hier lädt auch WL (-.63) und m (.58).

Free Recall lädt nicht in Faktor 1 aber hoch in Faktor 2 (.78).

Bildhaftigkeit-Konkretheit Variablen können als Vorhersage der Free Recall Leistung angesehen werden. FRINCKE nimmt an, daß bildhafte Wörter leichter in Cluster zu gruppieren sind, aber auch daß die Vpn mehr Aufmerksamkeit auf diese Wörter lenken.

Als Ergebnis kann also abschließend festgehalten werden, daß der Faktor 'Konkretheit-Bildhaftigkeit' beim Lernen eine wichtigere Rolle zu spielen scheint als die Bedeutungshaltigkeit und die Vorkommenshäufigkeit.

In den beiden Untersuchungen (PAIVIO 1968, FRINCKE 1968) bestand keine signifikante Beziehung zwischen der angenehmen Qualität (G) von Wörtern und dem Free Recall-Lernen*, ebenso keine zwischen der Intensität der emotionalen Tönung (E und e-p) und dem Free Recall-Lernen**. Man könnte also vermuten, daß die Emotionalität eines Wortes für das Lernen (Free Recall) keine Rolle spielt. Ein solcher Schluß wäre aus diesen Daten aber nicht erlaubt. Die Lernlisten vieler Untersuchungen sind nicht optimal konstruiert. Nur eine Lernliste, in der die neutralen und emotionalen Wörter bzgl. der anderen Wortparameter parallelisiert sind, erlaubt Rückschlüsse auf den möglichen Einfluß des emotionalen Erregungspotentials.

Ein weiterer wichtiger Punkt ist die Operationalisierung der 'Emotionalität'. In den beiden oben erwähnten Studien wurde die Emotionalität nur über das subjektive Urteil ermittelt, das nicht unbedingt sehr valide sein muß. Wie schon angeführt divergieren die subjektiven Einschätzungen und die physiologischen Reaktionen der Vpn. Es könnte sein, daß die Emotionalität keinen Einfluß hatte, weil sie durch die Urteile nicht richtig erfaßt wurde.

* $r_{G,FR} = -.19$ (PAIVIO); $r_{G,FR} = .14$ (.19/andere Variablen konstant)

** $r_{E,FR} = .02$ (PAIVIO); $r_{e-p,FR} = .17$ (.12/ " , FRINCKE)

4. Einfluß weiterer Faktoren auf das verbale Lernen

Das folgende Kapitel befaßt sich mit weiteren Einflußgrößen auf das verbale Lernen, wie andere Parameter des Lernstoffs oder Situationseinflüsse, die in der späteren Untersuchung beachtet werden müssen.

4.1. Charakteristika der Wortliste

4.1.1. Beziehungen zwischen den Items

Ein wichtiger Faktor, der die Gedächtnisleistung beeinflusst, ist die Art der Beziehung der Items untereinander. Gibt man Wörter vor, die eine enge assoziative Verbindung aufweisen und/oder kategorial zusammenhängen, zeigt sich, daß mehr Wörter erinnert werden als bei einer Liste, die aus unzusammenhängenden Wörtern besteht.

Die Untersuchung von DEESE (1959, zit. nach COFER 1971) erbrachte einen Korrelationskoeffizienten von .88 zwischen der 'assoziativen Stärke der Items zueinander' (inter-item associative strength, IIAS)* und der Anzahl der erinnerten Wörter.

Bei MURRAY (1974) war die Behaltensleistung ebenfalls größer bei Listen mit hoher IIAS als bei Listen mit niedriger IIAS oder keiner IIAS.

Diese Erleichterung des Lernens findet sich vor allem beim Free Recall. Beim seriellen Lernen, wo die richtige Reihenfolge verlangt wird, kann die Ähnlichkeit der Items auch erschwerend wirken (POSTMAN 1967, zit. nach HALL 1971).

Die Zusammensetzung der Wortliste aus kategorial verbundenen Wörtern verbessert die Behaltensleistung gegenüber unverbundenen Wörtern. (UNDERWOOD 1964, zit. nach COFER 1971; PUFF

* IIAS ist definiert als die durchschnittliche relative Häufigkeit, mit der alle Items in einer Liste alle anderen Items in der gleichen Liste als freie Assoziationen hervorrufen. (HALL 1971)

1970). Es ist günstig, wenn die Liste nicht nur aus einer einzigen Kategorie besteht, sondern aus mehreren, die jeweils ca. 5 Wörter umfassen.

Die Lernerleichterung wurde unter anderem mit Hilfe von 'impliziten Assoziationen' (implicit associational response, IAR) erklärt. (LEICHT 1968) Assoziativ verbundene Wörter werden besser erinnert als unverbundene, weil sie zusätzlich zu ihrer Präsentation als IARs zueinander auftreten, dadurch erfahren die Wörter eine implizite Wiederholung. Die Wirkung geht auf die Häufigkeit zurück, mit der ein Wort in einer Liste auftaucht. LEICHT konnte in seiner Untersuchung zeigen, daß die Reproduktionswahrscheinlichkeit und die beurteilte Häufigkeit zunahmen mit der Anzahl von Verbindungen eines Wortes mit anderen Listenwörtern.

Neben der Zunahme behaltener Items ist das sogenannte 'Clustering' ein weiterer beobachtbarer Effekt. Wörter, die näher verwandt sind, werden bei der Reproduktion unmittelbar nacheinander (in Form eines Clusters) reproduziert, auch wenn sie bei der Darbietung durch andere Wörter voneinander getrennt waren. Die Vpn erinnern die Wörter nicht in zufälliger Reihenfolge, sondern bilden Gruppen (Cluster) zusammengehöriger Wörter. Der Hinweis, daß die Items in Kategorien vorliegen, kann den Clustereffekt erhöhen, vor allem wenn die Items öfters wiederholt werden (JEFFMAR 1975). Ebenso erleichtert die Präsentation der Wörter in Blöcken (die Wörter einer Kategorie erscheinen hintereinander und nicht über die Liste verstreut) die Lernleistung. Die erste Studie zum kategorialen Clustering stammt von BOUSFIELD (1953, zit. nach HALL 1971). Es wurden verschiedene Formeln erstellt, um die Stärke des Clustering quantitativ zu erfassen. MARX (1980) entwickelte eine Methode, die eine geometrische Darstellung des Clustering erlaubt. Die Organisationsstruktur wird dabei sichtbar gemacht. Wörter, die nahe beieinanderliegend reproduziert werden, liegen auch in der Abbildung nahe beieinander.

In früheren Untersuchungen wurde häufig der Fehler begangen, emotionale Wörter auszuwählen, die nur aus wenigen Kategorien stammten, während die neutralen Wörter willkürlich zusammengestellt wurden. Dieser Fehler unterlief z.B. MANNING & MELCHIORI (1974) u.a. (siehe dazu Kritik bei MANNING & JULIAN 1975). Die Überlegenheit der emotionalen Items könnte durch diesen Faktor verursacht worden sein, da Wörter, die enger zusammenhängen, bei einer anschließenden Behaltensprüfung leichter erinnert werden.

4.1.2. Position

Die Stelle, an der ein Item in der Liste erscheint, spielt beim Lernen eine Rolle. Bei der Behaltensprüfung unmittelbar im Anschluß an die Darbietung der Stimuli werden primär die Items am Ende der Liste erinnert. Dies gilt für die freie Reproduktion wie für das Paarassoziationslernen. In der Literatur wird dieses Phänomen als 'Recency-Effekt' bezeichnet, da das zeitlich nächstliegende Material bevorzugt reproduziert wird. Der Recency-Effekt zeigt sich auch bei einer etwas verzögerten Reproduktion (ca. 30 - 60 sec später), wenn in der Zwischenzeit keine andere Aktivität stattfindet (BADDELEY & HITCH 1976, zit. nach BADDELEY 1979). Wird aber eine andere Aufgabe eingeschoben (meistens werden Zähl- oder leichte Rechenaufgaben verwendet), verschwindet der Recency-Effekt, die letzten Items einer Liste werden nicht mehr besser behalten (POSTMAN & PHILLIPS 1965, zit. nach BADDELEY; GLANZER & CUNITZ 1966, zit. nach GRUNEBERG 1972). Nach einem längeren Retentionsintervall werden diese Wörter sogar am seltensten erinnert (CRAIK 1970, MADIGAN & McCABE 1971; zit. nach M.W.EYSENCK 1976).

Eine weitere Sonderstellung nehmen die Items am Anfang einer Liste ein. (Primacy-Effekt) Auch sie werden besser behalten als die Items in der Mitte einer Lernliste. Beim seriellen Lernen werden sie schneller gelernt. Beim Free Recall werden unter der Bedingung des absichtlichen Lernens vor allem die

Wörter in der Anfangsposition reproduziert (SCHÖNPFLUG 1963).

Die verschiedenen Free Recall Studien, bei denen Listenlänge (10 - 40 Items) und Darbietungszeit variiert wurden, erbrachten immer, daß die ersten 3 bis 4 Items und die letzten 5 bis 7 Items besser behalten werden als die mittleren Items (z.B. DEESE & KAUFMAN 1957, zit. nach HALL 1971; MURDOCK 1962, zit. nach KINTSCH 1977).

SCHÖNPFLUG (1963) nimmt an, daß der gleiche Lernstoff an unterschiedlichen Stellen einer Lernliste unterschiedlich aktivierend ist. Am Anfang der Liste zeigt sich eine starke elektrodermale Aktivität, die im Verlauf der Listenpräsentation kontinuierlich abnimmt (SCHÖNPFLUG & BEIKE 1964, WALKER & TARTE 1963). Bei WALKER & TARTE ähnelte bei der Langzeitretention (1 Woche) die Verteilung des Recall den Hautwiderstandsänderungen. Vor allem die ersten Items, bei denen die EDR am größten waren, wurden erinnert.

C.H.BROWN (1937), der das serielle Lernen 19 sinnloser Silben mit der Antizipationsmethode untersuchte und dabei die EDR aufzeichnete, erhielt eine Rangkorrelation zwischen EDR der verschiedenen Itempositionen und Lernleistung der Itempositionen von .80. Am Anfang und am Ende der Liste fielen die EDR höher aus. Nach BROWNS Ansicht widmen die Vpn den Items am Anfang und am Schluß größere Aufmerksamkeit.

4.2. Situationseinflüsse

4.2.1. Art der Darbietung

Die Behaltensleistung von Wörtern kann in einem Lernversuch durch die Art (akustisch, optisch) und durch die Länge der Darbietung beeinflusst werden.

Bei der akustischen Darbietung der Wörter ist der Recency-Effekt ausgeprägter als bei der visuellen Präsentation (MURDOCK & WALKER 1969, zit. nach BADDELEY 1979).

UEHLING (1972) vermutete, daß die akustische Präsentation mit verantwortlich dafür war, daß manche Studien nicht das von WALKER postulierte Ergebnis erbrachten, sondern auch bei der Kurzzeitgedächtnisprüfung eine Überlegenheit der stark erregenden Items aufwiesen.*

Bei einer längeren Darbietungszeit (mehrere Sekunden pro Item) können mehr Items behalten werden. MANNING & GOLDSTEIN (1976) erhielten bei einer 4 sec Präsentationsrate eine signifikant größere Reproduktionsmenge und ein ausgeprägteres Clustering als bei einer 1 sec Darbietungsrate.

Bei zeitlich kürzerer Wortfolge (1 sec) waren Tabuwörter den neutralen Wörtern in der Behaltensprüfung deutlicher überlegen als bei längerer Präsentationsrate (4 sec). (MANNING & GOLDSTEIN 1976). Dieses statistisch nicht signifikante Ergebnis steht im Widerspruch zu der Aussage von KAPLAN & KAPLAN (1968, in ihrer Kritik an der Untersuchung von YARMEY 1966), daß sich bei einer zu schnellen Präsentation (2 sec) kein Arousal etablieren könne, da die EDR zwei Sekunden bis zu ihrem Auftreten braucht.**

4.2.2. Absichtlichkeit vs. Unabsichtlichkeit des Lernens

Auch wenn nicht absichtlich gelernt wird, bleibt einiges im Gedächtnis haften. Daß trotz fehlender Lernabsicht ein Einprägen von Wahrgenommenem stattfindet, wurde schon früh er-

* Aus heutiger Sicht läßt sich diese Aussage nicht mehr aufrechterhalten, da auch bei visueller Darbietung manchmal stark erregende Items kurzfristig besser behalten wurden.

** Das Arousal kommt auch bei einer kurzen Darbietungsfolge zum tragen. Der Zeitpunkt des Erscheinens der EDR ist nicht identisch mit dem Zeitpunkt der Auslösung. Allerdings dauert es etwa zwei Sekunden bis die EDR bei der Messung sichtbar wird. Ein gewisser Nachteil bei einer schnellen Reizdarbietungsfolge ist, daß sich die EDR-Ausschläge überlagern können, wenn eine neue Erregung ausgelöst wird, bevor die vorhergehende abgeklungen ist.

kannt (RADORSOWLJEVITCH 1907, POPPELREUTER 1912, MYERS 1913; zit. nach SCHÖNPFLUG 1963).

Meist werden aber unter der Bedingung des absichtlichen Lernens mehr Items behalten. Intentionales Lernen führt im allgemeinen zu einer besseren Reproduktionsleistung als inzidentelles Lernen.

SCHÖNPFLUG (1963 und 1964 zusammen mit BEIKE), der untersuchte, ob die 'Aktivation' (Erregungszustand) mit verschiedenen Lernbedingungen variiert, wies nach, daß durch die Lernintention die Aktivierung erhöht wurde (sowohl bzgl. des Basalwiderstandes als auch bzgl. der Amplitudenwerte). Eine bessere Lernleistung war die Folge.

In der Literatur taucht oft die Vermutung auf, daß der Einfluß der affektiven Tönung des Lernmaterials vor allem unter der Bedingung des unabsichtlichen Lernens zutage tritt (RAPAPORT 1977, ANISFELD & LAMBERT 1966). ANISFELD & LAMBERT glauben, daß der Effekt beim unabsichtlichen Lernen größer ist: "affective tone could be expected to show its maximum effects ... when motivational factors urging mastery of all items are minimized through incidental learning instructions." In der Untersuchung von AMSTER zeigte sich aber, daß beim unabsichtlichen Lernen gegenüber dem absichtlichen Lernen mehr neutrale Wörter erinnert wurden.

Ob also der Effekt beim unabsichtlichen Lernen oder beim absichtlichen Lernen stärker ist oder ob kein Unterschied besteht, konnte bisher noch nicht geklärt werden. Die Untersuchungen zum Einfluß der emotionalen Tönung verwendeten meist nur eine Form des Lernens, entweder unabsichtliches oder absichtliches. Seltener sind Untersuchungen, die beide Arten heranzogen, aus ihnen lassen sich jedoch keine eindeutigen Schlüsse ziehen. Aus den Daten von SCHÖNPFLUG & BEIKE (1964) läßt sich eine leichte relative Überlegenheit des emotionalen Lernmaterials beim unabsichtlichen Lernen ermitteln. Die Differenz zwischen stark ich-bezogenem Stoff und schwach ich-bezogenem Material ist beim inzidentellen Lernen größer.

4.2.3. Länge des Retentionsintervalls

Man kann in der Forschung drei Arten von Retentionsintervallen unterscheiden: die sofortige, die den Ultrakurzzeitspeicher betrifft, die Kurzzeitbehaltensprüfung (short term, ST) und die Langzeitbehaltensprüfung (long term, LT).

M.W.EYSENCK (1976) und LEVONIAN (1972) subsummieren unter die Kurzzeitbehaltensprüfungen solche, die unmittelbar im Anschluß an die Darbietung oder ein paar Minuten später beginnen. Reproduktionsprüfungen, die erst nach mehreren Minuten (laut EYSENCK 15 bis 20 Minuten, laut LEVONIAN 30 Minuten), Stunden, Tage oder Wochen erfolgen, fallen unter die Langzeitbehaltensprüfungen.

Unmittelbar im Anschluß an die Darbietung des Lernstoffs ist die Behaltensleistung am besten, mit länger werdendem Intervall verschlechtert sich die Leistung immer mehr.

Schon früh wurde die Ansicht vertreten, daß sich die Wirkung des emotionalen Faktors erst bei einer späteren Behaltensprüfung bemerkbar macht. Bei Kurzzeitgedächtnisprüfungen wird sein Einfluß als eher unbedeutend angesehen. (PETERS 1911, WHITE & RATLIFF 1934, GILBERT 1977; zit. nach RAPAPORT 1977)

In der Theorie von WALKER & TARTE (1963) wird eine Interaktion zwischen Länge des Retentionsintervalls und Arousal postuliert. Demnach sollen beim kurzfristigen Behalten die erregenden Items schlechter reproduziert werden, beim langfristigen Behalten dagegen besser.

LEVONIAN (1972) vermutet, daß beim Kurzzeitgedächtnis zwischen Arousal und Gedächtnisleistung eine umgekehrte U-Funktion besteht, beim Langzeitgedächtnis aber eine positiv lineare.

Bei CORTEEN (1969) wurde mit zunehmendem Retentionsintervall die Beziehung zwischen EDR und Behaltensleistung größer. In der Untersuchung von WARREN & HARRIS (1975) wurden dagegen nur bei der Kurzzeitgedächtnisprüfung mehr erregen-

de als neutrale Items erinnert. Die Langzeitgedächtnisprüfung ergab keinen Unterschied in der Behaltensleistung von neutralen und erregenden Items, die neutralen Wörter wurden gleich gut reproduziert.

4.3. Erregung und verbales Lernen

Abschließend sei noch eine Reihe von Untersuchungen erwähnt, auf die in dieser Arbeit nicht näher eingegangen wird. Auch sie untersuchen die Beziehung zwischen Erregung und verbalem Lernen, allerdings wird die Erregung nicht durch die Art des Lernmaterials hervorgerufen.

Drei Gruppen lassen sich unterscheiden:

- 1) Die Erregung wird künstlich herbeigeführt durch Gabe von Drogen, verzögerte akustische Rückmeldung, induzierte Muskelspannung oder weißes Rauschen (gebräuchlichste Methode). BERLYNE et al. (1965, 1968), McLEAN (1969) und WESNER (1972) verwendeten induzierte Erregung (weißes Rauschen) während des Lernens. HÖRMANN & OSTERKAMP (1966), UEHLING & SPRINKLE (1968) und M.W.EYSENCK (1975) benutzten weißes Rauschen während der Gedächtnisprüfung.
- 2) Die Erregung wird durch 'Anreize' (Geld, Ankündigung eines E-Schocks) erzeugt. (HEINRICH 1968, PALLAK & ANDREW 1976, LUSTIG & WITRYOL 1977)
- 3) Die Erregung entsteht nicht durch direkte Manipulation, statt dessen wird die vorhandene Grunderregung der Vp herangezogen. Die Feststellung der Erregung geschieht meist durch eine der folgenden Möglichkeiten:
 - durch physiologische Messung (C.H.BROWN 1937; BERRY 1962; KLEINSMITH, KAPLAN & TARTE 1963; EISDORFER 1968)
 - durch subjektive Einschätzung der Erregung ('Activation-Deactivation Adjective Check List' von THAYER 1967; KATAHN, SYNDER & DURLACK 1970)
 - durch Erschließen über Persönlichkeitsfaktoren (Introversion - Extraversion; Neurotizismus; Ängstlichkeit) (HOWARTH & H.J.EYSENCK 1968; AMELANG, WENDT & FRÜNDT 1977)

Diese Untersuchungen lieferten unterschiedliche Ergebnisse: manche weisen auf eine positiv lineare Beziehung zwischen Erregung und Behalten hin, andere tendieren zu einer umgekehrten U-Funktion und einige entsprechen der Annahme von WALKER & TARTE (1963).

II EMPIRISCHER TEIL

1. Fragestellungen und Hypothesen

1.1. Emotionales Erregungspotential

In dieser Arbeit soll der Einfluß des emotionalen Erregungspotentials von Wörtern auf das verbale Lernen beim Free Recall untersucht werden.

Die zentrale Frage lautet: Werden emotional erregende* Wörter besser behalten als neutrale Wörter.

Hypothese 1: Emotional erregende Wörter bleiben besser im Gedächtnis haften als neutrale Wörter. Ereignisse, die eine emotionale Erregung auslösen, sind für den Organismus von größerer Bedeutung als neutrale Ereignisse und werden deshalb besser gemerkt.

H_0 : Neutrale Wörter und emotional erregende Wörter werden gleich gut behalten.

H_1 : Die Behaltensleistung bzgl. emotional erregender Wörter ist größer als bzgl. neutraler Wörter.

Die Operationalisierung des emotionalen Erregungspotentials erfolgt auf drei verschiedene Arten:

- 1) Die Emotionalität eines Wortes wird aufgrund von Befragungen an einer größeren Stichprobe gewonnen. (Normative Bestimmung)
- 2) Die durch ein Wort hervorgerufene emotionale Erregung wird physiologisch anhand der Hautleitfähigkeitsreaktion gemessen. (Ipsative Bestimmung)
- 3) Die empfundene emotionale Erregung wird durch subjektive Einschätzung auf einer Ratingskala erhoben. (Ipsative Bestimmung)

* 'Emotional erregend' bedeutet hier: Die Erregung hängt nur vom Wort selbst ab und nicht von dessen Stellung in der Wortliste.

Bei 2) und 3) geschieht die Auswertung ipsativ, da davon ausgegangen wird, daß die emotionale Bedeutung eines Wortes interindividuell stark variieren kann. Vermutet wird, daß bei einer ipsativen Betrachtungsweise der Effekt des emotionalen Erregungspotentials deutlicher sichtbar wird als bei einer normativen Auswertung.

Hypothese 2:

H₀: Es besteht kein Unterschied im Anteil emotional erregender Wörter an der Behaltensleistung, gleichgültig ob die emotionale Erregung normativ oder ipsativ bestimmt wird.

H₁: Bei der ipsativen Methode fällt das Verhältnis von emotionalen behaltene[n] Wörtern zu neutralen behaltene[n] Wörtern größer aus.

Weiter soll geprüft werden, welche Rolle die physiologische Reaktion (2) und welche Rolle die subjektive Empfindung (3) spielt. Wird die Gedächtnisleistung mehr von den unbewußten physiologischen Vorgängen oder von der bewußten subjektiven Empfindung bestimmt?

Hypothese 3: Erwartet wird (aufgrund von Vorversuchen), daß bei Zugrundelegung der physiologischen Werte das Verhältnis von emotionalen behaltene[n] Wörtern zu neutralen behaltene[n] Wörtern größer ist als bei den auf subjektivem Empfinden basierenden Werten.

H₀: Es besteht kein Unterschied im Anteil emotional erregender Wörter an der Behaltensleistung unabhängig davon ob die physiologischen Werte oder die subjektiv eingeschätzten Werte verwendet werden.

H₁: Der Anteil emotional erregender Wörter an der Behaltensleistung ist nach der Klassifikation der Wörter aufgrund der physiologischen Reaktion größer.

Der Einfluß des emotionalen Erregungspotentials von Wörtern soll unter verschiedenen Bedingungen analysiert werden:

- a) absichtliches und unabsichtliches Lernen
- b) kategoriale Wortliste und Liste aus unverbundenen Wörtern

- c) Kurzzeitbehaltensintervall und Langzeitbehaltensintervall
- d) Geschlecht.

Es soll untersucht werden, ob diese Variablen die Wirkung der emotionalen Erregung beeinflussen. Tritt der Einfluß des emotionalen Erregungspotentials unterschiedlich stark in Erscheinung oder bleibt seine Wirkung unverändert, gleichgültig unter welchen Bedingungen das Lernen stattfindet?

1.2. Emotionales Erregungspotential und Absichtlichkeit des Lernens (Intentionales vs. inzidentelles Lernen)

Es wird angenommen, daß sich der Effekt der emotionalen Erregung beim unabsichtlichen Lernen stärker bemerkbar macht als beim absichtlichen. Beim intentionalen Lernen ist die Vp bemüht, sich alle Wörter zu merken und wendet daher Einprägungsstrategien an, die ihr helfen, auch neutrale Wörter zu behalten. Vermutlich ist auch im Lernexperiment die Haltung den Wörtern gegenüber gleichmäßiger, so daß Unterschiede in emotionalen Reaktionen auf die einzelnen Wörter nicht so stark auftreten wie beim inzidentellen Lernen. Beim unabsichtlichen Lernen soll die Vp nur die Wörter lesen und auf sich wirken lassen. In dieser Situation kommt es möglicherweise zu stärkeren emotionalen Reaktionen bei den für die Vp emotional behafteten Wörtern.

Es geht hier nur um spezifische emotionale Reaktionen, die durch die Wörter ausgelöst werden, und nicht um situationsbedingte Erregungen. Ob z.B. die Lern- oder die Lesesituation als belastender empfunden wird, hängt von vielen Faktoren ab. Für einige Vpn mag die Lernsituation, bei der es um eine Leistung geht, stärker erregend wirken, für andere die Lesesituation, da hier die physiologische Messung mehr in den Blickpunkt rückt und sich die Vp zu sehr beobachtet fühlen kann.

Hypothese 4: Der Einfluß des emotionalen Erregungspotentials von Wörtern ist beim unabsichtlichen Lernen größer als beim absichtlichen Lernen.

H_0 : Es besteht kein Unterschied im Anteil emotional erregender Wörter an der Behaltensleistung zwischen absichtlichem und unabsichtlichem Lernen.

H_1 : Beim unabsichtlichen Lernen werden prozentual mehr emotional erregende Wörter erinnert.

1.3. Emotionales Erregungspotential und Zusammensetzung der Lernliste (Clusterliste vs. Willkürliste)

Unter 'Clusterliste' wird hier eine Liste verstanden, die aus Gruppen inhaltlich zusammengehöriger Wörter besteht, die innerhalb einer Kategorie starke Interitemassoziationen besitzen.

Eine 'Willkürliste' setzt sich aus Wörtern zusammen, die untereinander keine Beziehungen aufweisen, also durch geringe Interitemassoziationen gekennzeichnet sind.

Hypothese 5: Es wird vermutet, daß bei der Willkürliste, bei der eine Störung durch Assoziationen zwischen den Items vermieden werden soll, der Erregungseffekt stärker hervortritt als bei der Clusterliste. Die Clusterliste könnte die Vpn veranlassen, das Lernmaterial eher kognitiv als emotional zu verarbeiten.

H_0 : Es werden prozentual gleich viele emotional erregende Wörter von der Clusterliste wie von der Willkürliste behalten.

H_1 : Von der Willkürliste werden prozentual mehr emotional erregende Wörter erinnert.

1.4. Emotionales Erregungspotential und Länge des Retentionsintervalls (Kurzzeit- vs. Langzeitgedächtnis, ST - LT)

In den Untersuchungen zum Paarassoziationslernen zeigte sich bei der Prüfung des Langzeitgedächtnisses ein größerer Effekt der emotionalen Erregung.

Hypothese 6: Es wird erwartet, daß sich dieser stärkere Effekt auch unter der Free Recall Bedingung ergibt.

Wenn unmittelbar nach der Darbietung das Gedächtnis geprüft wird, besteht die Gefahr, daß vor allem die letzten und eventuell die ersten Wörter reproduziert werden. Durch diesen Positionseffekt kann die Wirkung der emotionalen Erregung überdeckt werden.

Kurzfristig bleibt auch das neutrale Material im Gedächtnis, langfristig werden vor allem die emotional erregenden Wörter erinnert.

H_0 : Zwischen Kurzzeit- und Langzeitgedächtnisleistung besteht kein Unterschied im prozentualen Anteil an emotional erregenden Wörtern.

H_1 : Bei der Langzeitgedächtnisprüfung werden prozentual mehr emotional erregende Wörter erinnert als bei der Kurzzeitgedächtnisprüfung.

Hypothese 7: Wird bei der Kurzzeitgedächtnisprüfung der Recency-Effekt durch eine Zwischenaufgabe (30 sec Zahlen subtrahieren) ausgeschaltet, sollte der Erregungseffekt stärker ausgeprägt sein.

H_0 : Der Recency-Effekt hat keinen Einfluß auf die Behaltensleistung bzgl. emotional erregenden Wörter.

H_1 : Bei Vermeidung des Recency-Effekts werden prozentual mehr emotional erregende Wörter behalten als ohne Kontrolle.

Die Störaufgabe wird nur bei der Willkürliste gegeben, da erwartet wird, daß bei der Clusterliste durch die Listenstruktur kein Recency-Effekt auftritt.

Beim Langzeitgedächtnis soll zusätzlich die Rekognition geprüft werden.

1.5. Emotionales Erregungspotential und Geschlecht

In der Literatur fand sich keine Untersuchung, bei der die Behaltensleistung von Männern und Frauen hinsichtlich der emotionalen Tönung des Lernmaterials verglichen wurde. Es lassen sich also nur Mutmaßungen darüber anstellen, ob Männer und Frauen unterschiedlich auf den emotionalen Erregungsfaktor reagieren.

Hypothese 8: Es besteht ein Unterschied im Behalten emotional getönter Wörter zwischen Männern und Frauen.

Die unterschiedlichen gesellschaftlichen Rollenbilder und Sozialisationserfahrungen von Mann (Verstandesbetonung) und Frau (Gefühlsbetonung) ließen erwarten, daß Männer weniger von der emotionalen Erregung beeinflusst werden.

Andererseits zeigen Untersuchungen, daß Jungen im Kindesalter vermehrt emotionale Reaktionen aufweisen als Mädchen. (HURLOCK 1970)

H_0 : Frauen und Männer unterscheiden sich nicht bzgl. des Anteils emotional erregender Wörter an der Gesamtgedächtnisleistung.

H_1 : Frauen erinnern prozentual mehr emotional erregende Wörter als Männer.

H_2 : Männer erinnern prozentual mehr emotional erregende Wörter als Frauen.

2. Operationalisierung der Variablen, Versuchsplan

2.1. Unabhängige Variablen

Emotionales Erregungspotential: Ein Wort ist

- normativ emotional erregend, wenn es bei einer Worterhebung als solches genannt wird und von einer Gruppe von Vpn als emotional erregend beurteilt wird.
- physiologisch emotional erregend, wenn auf das Wort eine EDR innerhalb einer Latenzzeit von 1 bis 3 Sekunden erfolgt.
- subjektiv emotional erregend, wenn die Vp auf einer graphischen Skala mit den Polen 'neutral' und 'maximal erregend' nicht den neutralen Pol kennzeichnet.

Es sollen zwei Stufen dieser Variablen herangezogen werden:

1) emotional erregende Wörter, 2) nicht emotional erregende Wörter = neutrale Wörter.

Es ist vorgesehen, zusätzlich eine Analyse der Daten nach der unterschiedlichen Intensität des emotionalen Erregungspotentials vorzunehmen, nach folgenden drei Erregungsgraden:

1) stark 2) mittel 3) schwach.

Es gibt kein Kriterium dafür, wann ein Wort stark, mittel oder schwach erregend ist. In dieser Arbeit fiel die Entscheidung nicht für absolute Werte (z.B. EDR-Amplitude von einer bestimmten Größe), sondern für relative Werte. Der jeweils höchste verwertbare Ausschlag der Vp wird als ihre maximale Erregung in diesem Versuch angesehen, dessen Höhe den Richtwert für die Einteilung in die drei Bereiche bildet. Er wird gleich 100% gesetzt. In die stark erregende Gruppe fallen Wörter mit einer Amplitude größer 75% des maximalen Ausschlags. Als mittel erregend gelten Wörter mit einer Amplitude von 25% bis 75% des maximalen Ausschlags, als schwach erregend Wörter mit einer Amplitude kleiner 25%.

Von den physiologischen Reaktionen her erscheint diese Vorgehensweise sinnvoller. Erinnert sei an das Ausgangswertgesetz:

Die Größe der Reaktion hängt von der Grunderregung ab. Würde man von vorneherein einen bestimmten Wert (in Microsie mens) festsetzen, könnte es passieren, daß einige Vpn mit keiner ihrer Reaktionen darüber liegen, andere dagegen mit allen ihren Reaktionen über diesem Grenzwert liegen.

Hinsichtlich der subjektiven Einschätzung wird das gleiche Verfahren gewählt, um die gleiche Ausgangslage für den Vergleich zwischen subjektiver und physiologischer Methode zu erhalten. Die von der Vp markierte maximale Entfernung vom neutralen Pol ist gleich 100%.

stark erregend	mittel erregend	schwach erregend
> 75%	75% - 25%	< 25%

Die Art der Beziehung zwischen der Intensität der emotionalen Erregung und dem Behalten von Wörtern soll untersucht werden. Liegt eine positiv lineare Beziehung vor oder eine kurvilineare Beziehung in Form eines umgekehrten U ?

Absichtlichkeit des Lernens (2 Stufen)

Die beiden Stufen dieser Variablen werden durch die Instruktion erstellt. Beim absichtlichen Lernen ist es die Aufgabe der Vp, die Wörter zu lernen, beim unabsichtlichen Lernen soll die Vp die Wörter nur lesen und auf sich wirken lassen.

Zusammensetzung der Wortliste (2 Stufen)

Willkürliste: Die Liste setzt sich aus möglichst unterschiedlichen Wörtern zusammen. Kein Wort soll mit einem anderen Wort der Liste eine bedeutungsmäßige oder assoziative Beziehung aufweisen. Die Liste soll außerdem verschieden Wortarten enthalten (Substantive, Adjektive und Verben).

Clusterliste: Die Liste besteht aus Wörtern, die bestimmten Kategorien angehören. Die Wörter sollen in einem Assoziationsversuch gewonnen werden, die häufigsten Reaktionen sollen ausgewählt werden. Diese Liste soll ein Clustering erleichtern.

Länge des Retentionsintervalls (2 Stufen)

Die Kurzzeitgedächtnisprüfung erfolgt im Anschluß an die Darbietung der Wortliste, die Langzeitgedächtnisprüfung ca. eine Woche nach dem Lernversuch.

2.2. Abhängige Variable

Behaltensleistung:

Anzahl der reproduzierten Wörter,

Anzahl der wiedererkannten Wörter (zusätzlich bei LT).

2.3. Versuchsplan

Für die Prüfung des emotionalen Erregungspotentials wurde ein 'within subject design' gewählt. Jede Vp bekommt sowohl neutrale als auch emotional erregende Wörter dargeboten.

Die anderen Bedingungen wurden in einem 'between subject design' geprüft. Jede Vp wird nur in einer bestimmten Art der Lernbedingung (absichtlich oder unabsichtlich, Willkürliste oder Clusterliste, ST oder LT) untersucht.

Jeder Versuchsbedingung sollten gleich viele männliche wie weibliche Vpn nach dem Zufall zugeordnet werden. (etwa 10)

			emotional erregend	neutral
absichtlich	ST	Clusterliste F		
		M		
		Willkürliste F		
		M		
unabsichtlich	ST	Clusterliste F		
		M		
		Willkürliste F		
		(ohne Kontrolle) M		
	LT	Willkürliste F		
		(mit Kontrolle) M		
		Clusterliste F		
		M		

3. Versuchsmaterial

3.1. Lernmaterial

Die Auswahl der Wörter für die Lernlisten war einer der schwierigsten und wichtigsten Punkte dieser Untersuchung. Die Entscheidung, welches Wort emotional erregend und welches neutral ist, sollte nicht der Experimentator fällen; sie sollte vielmehr aus der Population (Studenten) stammen, mit der der Lernversuch durchgeführt wird.

Wie in Punkt 3 der Arbeit deutlich geworden ist, üben verschiedene andere Wortcharakteristika einen Einfluß auf das verbale Lernen aus. Diese Störvariablen müssen in der Untersuchung sorgfältig kontrolliert werden. Emotional erregende Wörter und neutrale Wörter sollten sich hinsichtlich dieser anderen Wortcharakteristika nicht unterscheiden. Deshalb war eine Parallelisierung der emotionalen und neutralen Wörter bzgl. der Vorkommenshäufigkeit, Abstraktheit und Wortlänge notwendig.

3.1.1. Worterhebungen*

Um festzustellen, welche Wörter von Studenten als neutral und welche als emotional erregend angesehen werden, wurden im Mai 1979 und im November/Dezember 1979 Worterhebungen an 144 Studenten der LMU in München vorgenommen. Auf einem Blatt sollten 20 neutrale (N) und 20 emotional behaftete (E) Wörter (auch Tabuwörter) niedergeschrieben werden. Die Befragung erfolgte anonym in Gruppen unterschiedlicher Größe. Die Vpn sollten nur ihr Alter und ihr Geschlecht angeben.

27 Vpn machten keine Angaben zu Geschlecht und Alter. Der Rest setzte sich aus 69 weiblichen und 48 männlichen Vpn zu-

* Hier handelt es sich eigentlich schon um eine Gedächtnisuntersuchung, denn die Vpn mußten die Wörter aus ihrem Gedächtnis auswählen.

sammen. Das Durchschnittsalter belief sich auf 22.7 Jahre, die Altersspanne reichte von 19 bis 36 Jahren.

Einzelnen Vpn fiel es schwer, neutrale Wörter zu finden. Sie äußerten, daß für sie jedes Wort eine emotionale Bedeutung habe. Andere wiederum hatten Schwierigkeiten emotional behaftete Wörter anzugeben. Etwa 30% der weiblichen Vpn notierten mehr emotionale Wörter, 40% der männlichen Vpn produzierten mehr neutrale Wörter.

Insgesamt wurden 5000 Wörter erfaßt, 2500 neutrale und 2500 emotional behaftete Wörter. Die Wörter wurden nach E und N getrennt alphabetisch aufgelistet, ihre Auftretenshäufigkeit wurde registriert. Die Liste der emotionalen Wörter enthält 1468 verschiedene Wörter (types), 181 Wörter mehr als die Liste der neutralen Wörter mit nur 1287 verschiedenen Wörtern. Der Grund dafür könnte sein, daß es mehr emotional behaftete Wörter gibt oder aber, daß die emotionale Bedeutung stärker gestreut ist (jede Vp empfindet andere Wörter als emotional, bzgl. der neutralen Wörter besteht ein stärkerer Konsens). 179 Wörter erscheinen in beiden Listen. Diese Wörter wurden von einigen Vpn als neutral, von anderen als emotional angesehen. Insgesamt liegen 2576 verschiedene Wörter vor (2755 - 179); die beiden Listen haben 6.9% der types gemeinsam.

Aus den Ergebnissen der Worterhebung werden in Tabelle 14 die 14 häufigsten Wörter der E- und N-Kategorie unter Angabe ihrer Auftretenshäufigkeit (F_{WE}) aufgeführt; bei den restlichen Häufigkeitsklassen wird nur die Anzahl der verschiedenen Wörter angegeben.

Ein interessanter Aspekt der Worterhebung war, daß mit dem Anfangsbuchstaben E, I, U und V deutlich mehr emotionale Wörter als neutrale genannt wurden. Die neutralen Wörter überwogen beim Anfangsbuchstaben B.

EMOTIONAL BEHAFTETE WÖRTER

F_{WE}	Anzahl	Wort
49	1	Liebe
25	1	Tod
24	1	Haß
19	1	Freund(e)
18	1	Sonne
17	2	Musik Mutter
16	1	Freude
14	2	Angst Vater
13	2	Scheiße Universität
12	2	Frau(en) Kind(er)
9	9	
8	9	
7	11	
6	11	
5	25	
4	35	
3	74	
2	174	
1	1106	
1468 types		

NEUTRALE WÖRTER

F_{WE}	Anzahl	Wort
38	1	Haus
32	1	Tisch
30	1	Stuhl
23	1	Auto
21	1	Papier
20	2	gehen Straße
18	1	Baum
17	1	Schuhe
16	1	Kugelschreiber/Kuli
15	2	Buch Tafel
13	2	Mantel Wort
12	3	
11	6	
10	4	
9	10	
8	13	
7	13	
6	16	
5	20	
4	36	
3	66	
2	158	
1	928	
1287 types		

Tabelle 14: Anzahl der types und ihre Auftretenshäufigkeit (F_{WE}) in der E- und N-Liste, jeweils mit Nennung der 14 häufigsten Wörter.

Diese beiden Listen wurden auf zwei verschiedene Arten weiterverarbeitet:

- a) Die beiden Listen (E und N) wurden mehreren Studenten vorgelegt, die die Wörter kennzeichnen sollten, die ihnen in der N-Liste am neutralsten und in der E-Liste am emotional erregendsten erschienen. Dies geschah anonym.
- b) Die genannten Wörter wurden in verschiedene inhaltliche Bereiche (Kategorien) zusammengefaßt, was sich als äußerst schwierig erwies, da kein optimales Kategoriensystem existiert. Manche Wörter kann man mehreren Bereichen zuordnen, andere lassen sich schlecht einer Gruppe zuweisen. Das Hauptproblem besteht darin, voneinander unabhängige und gleichwertige Bereiche zu finden. In der folgenden Tabelle 15 sind die wesentlichsten Kategorien zusammengestellt; die umfangreichsten stehen dabei am Anfang der Liste.

E-Bereiche		N-Bereiche	
nach Anzahl der types	nach Anzahl der token	nach Anzahl der types	nach Anzahl der token
Liebe/Sexualität	Liebe/Sexualität	Haus/Wohnen	Haus/Wohnen
Gewalt/Verbrechen	Gefühle	Verkehr/Fortbewegung	Verkehr/Fortbewegung
Medizin/Krankheit	Gewalt/Verbrechen	Material	Schreibzeug/Werkzeug
Gefühle	Medizin/Krankheit	Schreibzeug/Werkzeug	Kleidung
Tod	Tod	Haushaltsgegenstände	Sprache
Toilette/Körperausscheidung	Familie	Sprache	Medien
Religion	Toilette/Körperausscheidung	Kleidung	Material
Schimpfwörter	Religion	math., physik. Begriffe	math., physik. Begriffe
Familie	Schimpfwörter	Begriffe	Haushaltsgegenstände
		Medien	
Gemischte Bereiche (sowohl E als auch N)			
nach Anzahl der types u. nach Anzahl der token			
Natur (biologisch, geographisch, meteorologisch)			
Körperteile; Wissenschaft/Studium; Nahrung			
Kultur/Freizeit; Politik			

Tabelle 15: Wesentliche Bereiche der Worterhebung.

3.1.2. Lernlisten

Beide Listen (Willkür- und Clusterliste) sollten gleich viele Wörter enthalten. Um den Anfangspositionseffekt (sowohl hinsichtlich der EDR als auch des Lernens) auszuschalten, wurden 6 Übungswörter vorausgegeben, wobei die Vp aber nicht wußte, daß es sich dabei um Übungswörter handelt. Als Testwörter wurden 40 Wörter ausgewählt, wobei gleich viele E- wie N-Wörter berücksichtigt wurden. Für diese Anzahl sprachen zwei Gründe:

- a) Bei einer kürzeren Liste könnte es zu einem Deckeneffekt kommen. Unter der Bedingung 'absichtliches Lernen der Clusterliste' könnten ohne weiteres 20 Wörter reproduziert werden. Die Vp sollte sich höchstens die Hälfte der Wörter merken können, um eine maximale Differenzierung zwischen E und N zu ermöglichen.
- b) Auf eine längere Liste (mehr als 46 Wörter) wurde wegen möglicher Ermüdungserscheinungen der Vpn verzichtet. (Hinweis aus der Voruntersuchung)

Die 6 Übungswörter sollten in keinem inhaltlichen Zusammenhang mit den Testwörtern stehen und auch untereinander keine Beziehungen aufweisen. Sie spiegeln nur formal die Palette der nachfolgenden Wörter wider, was die Wortlänge, die Abstraktheit und die Wortform betrifft.

3.1.2.1. Willkürliste

Die Wörter dieser Liste sollten untereinander nicht zusammenhängen. Aus dem Pool der als besonders emotional erregend beurteilten Wörter wurde je ein Wort aus je einem der verschiedenen Bereiche genommen, das am besten mit einem als besonders neutral gekennzeichneten Wort in folgenden Kriterien übereinstimmte: 1) Auftretenshäufigkeit in der Worterhebung, 2) Wortlänge, 3) Wortform und 4) geschätzte Konkret-Abstraktheit. Gleichzeitig mußte aber darauf geachtet werden, daß die N-Wörter inhaltlich weder untereinander noch

mit irgendeinem E-Wort verbunden sind. Die Wörter in der engeren Auswahl wurden von einer kleinen Gruppe von Vpn bzgl. ihrer Konkretheit oder Abstraktheit auf einer 5-Punkte Skala (A) mit den Polen 1 = 'konkret' und 5 = 'abstrakt' eingeschätzt.

Es stellte sich heraus, daß es unmöglich ist, eine Liste zusammenzustellen, die aus lauter unzusammenhängenden Wörtern besteht. Eine Verbindung zwischen verschiedenen Wörtern läßt sich immer herstellen. Auf alle Fälle sollte sich in der Willkürliste eine Beziehung nicht so leicht ergeben wie in der Clusterliste.

Manchmal war es nicht möglich ein in allen Kriterien übereinstimmendes Paar zu finden. In diesem Fall erfolgte dann ein Ausgleich mit einem anderen Paar.

Die Willkürliste umfaßt 18 E und N-Paare, und 4 gemischt beurteilte Wörter. In der folgenden Tabelle 16 sind die Testwörter nach ihrer Auftretenshäufigkeit in der Worterhebung (F_{WE}) geordnet aufgeführt, mit Angabe der beurteilten Konkretheit-Abstraktheit (A) und der Vorkommenshäufigkeit in der 'Deutschen Sprachstatistik' von MEIER (1967), $F(M)^*$. Am Ende der Tabelle stehen die verwendeten Übungswörter in der Reihenfolge ihrer Darbietung.

3.1.2.2. Clusterliste

Mehrere Bereiche wurden aus der E- und N-Liste ausgewählt, wobei sich immer ein E- und ein N-Bereich in ihrem type-token Umfang entsprachen. Berücksichtigt wurden solche Bereiche, die als besonders emotional erregend oder neutral beurteilt worden waren, und einige Bereiche, die E- und N-Wörter enthielten. Die Bezeichnungen der Kategorien wurden 15 Studenten vorgelegt, die 6 oder 7 Wörter darauf assoziieren sollten,

* Die Werte von MEIER werden als zusätzliche Vergleichsmöglichkeit angegeben. Für die Auswahl der Wörter spielten sie jedoch keine Rolle.

Tabelle 16:

WILLKÜRLISTE

<u>emotional behaftete Wörter</u>				<u>neutrale Wörter</u>			
Wort	F _{WE}	A	F(M)	Wort	F _{WE}	A	F(M)
Krieg	9	3.0	1281	Lampe	9	1.1	238
Prüfung	8	2.5	909	Kante	8	1.8	33
Streit	8	3.0	474	Luft	8	2.5	2054
Psychologie	6	4.1	35	Regenschirm	7	1.0	***
Kirche	6	2.7	1737	Treppe	6	1.2	281
Folter	6	2.2	41	Heizung	5	1.9	89
Orgasmus	4	1.7	-	Buchstabe	4	1.7	55
Sau	3	1.6	17	Rad	3	1.2	80
Geborgenheit	3	3.8	-*	Unterschied	2	3.8	524
Leiche	2	1.5	218	Stiefel	3	1.0	83
Bordell	2	2.2	-	Regal	2	1.6	-
Tränen	2	1.6	954	Sprache	2	3.4	2021
schwanger	2	1.4	13	rechteckig	1	1.6	-
Abgrund	1	2.7	151	Gegend	2	3.0	903
geisteskrank	1	3.4	30	ausgeglichen	1	2.4	89
kotzen	1	1.2	-	sehen	1	1.9	5583
Amputation	1	2.0	-	Abbildung	1	2.9	118
Verbrechen	1	3.2	460	Material	1	3.5	648

Wort	F _{WE} (E/N)	A	F(M)
Musik	17/2	3.4	837
Wohnung	5/5	2.3	685
Geld	8/1	1.5	1965
Ferien	4/1	2.4	84
Übungswörter:			
Farbe	2/8	2.6	733
neun	/2	3.4	264
Hand	/8	1.1	2412
schreiben	/6	2.4	1932
Begriff	/3	4.6	980
Verkehr	1/	2.9	1151

* geborgen 46

** Regen 500; Schirm 70

so daß zu jedem Bereich 100 Reaktionswörter vorlagen. Die Bereiche, die durch besonders hochfrequente Primärantworten gekennzeichnet waren oder fast nur aus Einzelnennungen bestanden, fielen weg. In die Clusterliste wurden nur jene Gruppen aufgenommen, bei denen ein E- und ein N-Bereich eine ähnliche Antwortverteilung zeigten. Wenn möglich sollten die drei häufigsten Assoziationen und der Kategorienname verwendet werden, unter der Bedingung, daß die Wörter sich auch in der Ursprungsliste befanden. Die Primärantworten konnten bei einigen Bereichen nicht verwendet werden, wenn sie nicht eindeutig waren (z.B. 'Ton'), mit einem anderen Wort aus einem anderen Bereich verbunden waren oder eine andere emotionale Qualität als beabsichtigt aufwiesen (Beispiel: 'Gemütlichkeit' zu 'Haus' wurde als nicht neutral angesehen).

E- und N-Wörter sollten wie bei der Willkürliste bzgl. Wortlänge, Abstraktheit und Häufigkeit (aus der Worterhebung und dem Assoziationsversuch) übereinstimmen.

Die Bereiche, bei denen zahlenmäßig die beste Übereinstimmung von E und N vorlag, wurden für die Liste ausgewählt. Falls keine in allen Punkten übereinstimmenden Paare gebildet werden konnten, wurden auch hier die nichtrealisierten Kriterien über ein anderes Paar ausgeglichen.

Die Clusterliste umfaßt 10 Kategorien à 4 Substantive. Bei vier Kategorien handelt es sich um reine E-Bereiche (Liebe, Tod, Toilette, Religion). In vier Kategorien befinden sich nur neutrale Wörter (Haus, Material, Haushaltsgegenstände, Sprache). Zwei Kategorien sind gemischt (Körperteile, Wissenschaft).

In der folgenden Tabelle 17 sind die Wörter angefangen von der umfangreichsten bis zur kleinsten Kategorie angeordnet, mit Angabe der Häufigkeit in der Worterhebung (F_{WE}), der Häufigkeit im Assoziationsversuch (F_{AS}), der beurteilten Konkretheit-Abstraktheit (A) und den Häufigkeitswerten von MEIER (1967). Die Tabelle enthält auch die Übungswörter.

Tabelle 17:

CLUSTERLISTE

emotional behaftete Wörter

Wort	F _{WE}	F _{AS}	A	F(M)
Liebe	49		3.4	6665
Zärtlichkeit	7	5	3.0	135
Kuß	6	4	1.8	233
Orgasmus	4	1	1.7	-
Tod	25		2.7	2043
Beerdigung	2	7	3.0	41
Sarg	3	5	1.2	90
Sterben	6	3	2.8	872
Klopapier	1	8	1.0	-
Gestank	1	7	1.4	-
Durchfall	1	4	1.2	-
Kot	1	1	1.0	24
Gott	7	6	4.6	4719
Kirche	6	6	2.7	1737
Hölle	1	2	4.2	321
Sünde	1	1	4.4	505

neutrale Wörter

Wort	F _{WE}	F _{AS}	A	F(M)
Haus	38		1.2	2412
Zimmer	6	5	1.2	1379
Schornstein	4	2	1.0	24
Dachstuhl	1	1	1.2	-*
Material	1		3.5	648
Eisen	3	8	2.0	439
Holz	4	8	2.3	584
Stoff	1	6	2.8	608
Gabel	4	3	1.0	117
Geschirr	1	2	1.8	42
Kanne	2	2	1.1	26
Teller	6	1	1.1	113
Wort	13	8	3.5	3770
Satz	6	6	3.7	694
Buchstabe	4	2	1.7	55
Begriff	3	1	4.6	980

Wort	F _{WE} (E/N)	F _{AS}	A	F(M)
Fuß	/2	10	1.1	1075
Hand	/8	6	1.1	6372
Busen	5/	3	1.1	307
Penis	5/	1	1.1	-
Studium	3/1	4	3.4	412
Universität	13/4	2	2.2	339
Prüfung	8/	1	2.5	909
Psychologie	6/	1	4.1	35

Übungswörter:

Gegenstand	/4		3.8	1018
Tier	2/7		1.6	376
Durchmesser	/1		4.4	174
Luft	/8		2.5	2054
Bleistift	/10		1.0	33
Straße	1/20		1.6	1025

Die beiden Listen entsprechen sich weitgehend in den Verteilungen der Auftretenshäufigkeit, der Wortlänge und der Abstraktheit. Die Mittelwerte befinden sich in Tabelle 18.

	Willkürliste	Clusterliste
F _{WE}	4.4	4.4*
Wortlänge (Buchstaben)	7.2	6.3
Abstraktheit	2.27	2.26

Tabelle 18: Mittelwerte von Auftretenshäufigkeit, Wortlänge und Abstraktheit der Willkür- und der Clusterliste.

Neun Wörter sind in beiden Listen enthalten, davon sechs als Testwörter (Buchstabe, Kirche, Material, Orgasmus, Prüfung, Psychologie) und drei abwechselnd als Übungs- oder Testwörter (Begriff, Hand; Luft).

Die Reihenfolge der Wörter wurde per Zufall bestimmt, mit der Einschränkung, daß inhaltlich ähnliche Wörter (betrifft vor allem die Clusterliste) nicht nahe beieinanderstehen. Ebenso sollten nicht mehrere emotional erregende Wörter hintereinander kommen. (Reihenfolge der Wörter siehe S.162) Um den Positionseffekt zumindest allgemein zu kontrollieren, erhielt die Hälfte der Vpn die Testwörter in umgekehrter Reihenfolge. Die 6 Übungswörter blieben gleich, nur die Testwörter änderten ihre Position.

3.2. Subjektive Einschätzskalen

Die Einschätzung der emotionalen Erregung wurde auf einer graphischen Ratingskala (siehe Abb.16) mit den Polen 'neutral' und 'maximal erregend' vorgenommen. Die Vpn sollten auf einem Kontinuum von 10 cm Länge ankreuzen, wie stark gefühlsmäßig erregend für sie das jeweilige Wort während des

* Die drei häufigsten Wörter (Liebe, Haus und Tod) wurden hier nicht miteinbezogen.

WILLKÜRLISTE

Farbe
neun
Hand
schreiben
Begriff
Verkehr

Tränen
Kante
Leiche
Buchstabe
ausgeglichen
sehen
Amputation
Musik
geisteskrank
Regal
Streit
Treppe
Kirche
Stiefel
Sau
Wohnung
Folter
Gegend
Orgasmus
Material
Luft
Geborgenheit
rechteckig
Krieg
Heizung
kotzen
Unterschied
Bordell
Rad
Psychologie
Regenschirm
schwanger
Ferien
Geld
Abgrund
Sprache
Verbrechen
Lampe
Prüfung
Abbildung

CLUSTERLISTE

Gegenstand
Tier
Durchmesser
Luft
Bleistift
Straße

Geschirr
Beerdigung
Hand
Prüfung
Haus
Kuß
Material
Kirche
Buchstabe
Tod
Gestank
Busen
Teller
Psychologie
Dachstuhl
Liebe
Begriff
Gott
Holz
Sterben
Gabel
Klopapier
Zimmer
Orgasmus
Satz
Kot
Sünde
Studium
Penis
Schornstein
Eisen
Sarg
Wort
Hölle
Durchfall
Kanne
Zärtlichkeit
Universität
Fuß
Stoff

Reihenfolge der Wörter in den beiden Lernlisten

Versuchs war. Die Wörter wurden der Vp in der gleichen Reihenfolge vorgelegt wie beim Lernversuch.



4. Versuchsdurchführung

4.1. Ort und Zeit

Die Versuche fanden im Psychologischen Institut der Universität München (Raum 480 und 487) in der Zeit von Januar 1980 bis Juni 1980 statt. Die Raumtemperatur war relativ konstant (um 20°C).

4.2. Versuchsbeschreibung

Die Vp saß möglichst entspannt in einem bequemen Sessel, den Atemgürtel in Höhe des Zwerchfells umgelegt. Zur Messung der Hautleitfähigkeit wurden drei Elektroden an der rechten Hand appliziert. Die Stellen, an denen die Elektroden angebracht wurden, der Daumenballen und der Kleinfingerballen (für die Messung) und der Handrücken (für die Erdung), wurden vorher mit etwas Alkohol gereinigt, um für einen besseren Halt der Kleberinge zu sorgen. Die Elektroden wurden sorgfältig mit Elektrodencreme gefüllt, so daß die Creme exakt und gleichmäßig bis zum Elektrodenrand reichte. Die Elektroden wurden dann mit Hilfe der Kleberinge auf der Handinnenfläche und dem Handrücken befestigt. Dabei mußte darauf geachtet werden, daß die Elektroden eben aufliegen und es zu einem vollständigen Kontakt mit der Hautoberfläche kommt. Hatte eine Vp feuchte Hände, mußte mit zusätzlichen Kleberingen oder einem Pflaster nachgeholfen werden. Die Hand mit den Elektroden sollte locker auf die Sessellehne aufgelegt werden. Wichtig war, daß die Hand möglichst ruhig lag und die Sessellehne nicht umklammerte, da Bewegungen oder unterschiedlicher Druck zu Meßstörungen führen.

Am Psychophysiographen, der auf einem Tisch hinter der Vp stand, wurden die zwei Einschübe (EDA und SPU) eingeschaltet und der für die Vp zutreffende Meßbereich eingestellt. Zur Überprüfung der richtigen Einstellung für die Atmungsableitung mußte die Vp tief einatmen. blieb die Nadel hierbei ge-

rade noch im Schreibbereich, war die Einstellung richtig. Die Größe der elektrodermalen Reaktion konnte dabei abgelesen werden, da tiefes Einatmen zu einer EDR führt. Die EDR ließ sich auch durch ein lautes Geräusch hervorrufen (der V1 klatscht in die Hände). Fiel die Amplitude sehr klein aus, mußte ein empfindlicherer Meßbereich gewählt werden, im umgekehrten Fall (der Zeigerausschlag war zu stark) ein unempfindlicherer. Die richtige Einstellung war oft ein Glücksspiel. Manche Vpn befanden sich mit ihren Werten genau zwischen zwei Meßbereichen. Von Nachteil war, daß sich der Meßbereich nicht kontinuierlich einstellen ließ, sondern nur feste Meßbereiche wählbar waren. Da nicht vorhersehbar war, wie eine Vp im folgenden Versuch reagierte (Aktivierungszunahme oder -abnahme) bestand die Gefahr, daß ein ungünstiger Meßbereich eingestellt wurde. Während des Versuchs mußte dann umgeschaltet werden, wodurch mindestens ein Item für die Messung verloren ging.

Nachdem die Vp an den Psychophysiographen angeschlossen war und die Einstellungen überprüft waren, wurde ihr die für die jeweilige Bedingung (absichtliches oder unabsichtliches Lernen) entsprechende Instruktion gegeben. Hatte die Vp die Instruktion verstanden und war sie bereit, wurde mit der Projektion der Wörter begonnen.

Die Wörter wurden der Vp per Dia auf einem grauen Projektionsschirm (60x60 cm groß) dargeboten, der einen Meter vor der Vp aufgestellt war. Die Buchstabenhöhe betrug 2 cm. Alle Wörter waren mit der gleichen Schreibmaschine getippt (Substantive mit großem, alle anderen Wortarten mit kleinem Anfangsbuchstaben) und anschließend fotografiert worden. Auf dem Schirm erschienen die Wörter in weißer Schrift.

In diesem Experiment fiel die Entscheidung für eine visuelle Darbietung der Wörter, um zusätzliche Einflüsse, die bei einer akustischen Darbietung auftreten, wie z.B. Wechsel in der Stimmlage, Aussprache, Lautstärke und stimmlichen Färbung auszuschließen, die vermutlich einen Einfluß auf die

EDR ausüben könnten. Diese Größen können sicher nicht konstant gehalten werden, da für den Sprecher die Wörter auch nicht neutral sind und er beim Aussprechen unbewußt seine Einstellung den Wörtern gegenüber mit ausdrückt.

Insgesamt wurden 46 Dias gezeigt, zuerst die 6 Übungswörter und dann die 40 Testwörter. Alle 6 Sekunden wurde vom V1, der sich hinter der Vp befand, durch Tastendruck der Diawechsel ausgelöst. Von dem zweipoligen Taster ging eine zweite Verbindung zum Psychophysiographen, so daß der Knopfdruck dort registriert wurde. Der Diawechsel dauerte 1 1/2 Sekunden. In dieser Zeit war der Projektionsschirm dunkel. Das Wort blieb 4 1/2 Sekunden sichtbar. Die Zeit von 6 Sekunden wurde gewählt, weil bei kürzerer Diafolge die Gefahr von Überlagerungen der SCRs bestand (Latenzzeit: 1 - 3 sec, Reaktionsdauer: 2 - 3 sec). Längere Präsentationsraten (8 - 10 sec) hatten in Vorversuchen Ermüdungs- und Ablenkungseffekte bei den Vpn bewirkt.

Nach dem letzten Diawechsel erschien durch das Licht des Projektors eine helle Fläche auf dem Projektionsschirm, die bei den Vpn eine SCR hervorrief, entweder wegen der Lichtblendung oder wegen des Wechsels in der Reizform. An dieser SCR konnte man die durchschnittliche Latenzzeit der Vp ablesen. Die Diaprojektion dauerte insgesamt 4 Minuten 40 Sekunden. Nach Beendigung der Darbietung wurden der Diaprojektor und die Einschübe des Psychophysiographen ausgeschaltet, von der Vp wurden die Elektroden und der Atemgürtel abgenommen und die Hand von der Elektrodencreme gereinigt. Sollte das Kurzzeitgedächtnis geprüft werden, bekam nun die Vp ein Blatt Papier und einen Kugelschreiber und sollte alle Wörter aufschreiben, an die sie sich noch erinnerte, so lange bis sie glaubte, daß ihr kein Wort mehr einfiel. Es wurden ihr etwa 5 Minuten Zeit gelassen.

Im Anschluß an die Gedächtnisprüfung erhielt die Vp die subjektive Einschätzungsskala, auf der sie anstreichen sollte, wie stark erregend sie die einzelnen Wörter empfand. Abschließend

wurden der Vp einige Explorationsfragen gestellt bzgl. ihrer Konzentration, Aufmerksamkeit, Befolgen der Instruktion und eventueller Störungen. Falls Interesse bestand, wurden der Vp die abgeleiteten Kurven erklärt.

Bei den meisten Versuchen waren zwei V1 anwesend; einer war für den Diawechsel zuständig, der andere für die Beobachtung der Vp und die Kontrolle des Psychophysigraphen. Die V1 befanden sich nicht im Sichtfeld der Vp. Der zweite V1 notierte Besonderheiten des Verhaltens der Vp (Husten, starke Bewegungen).

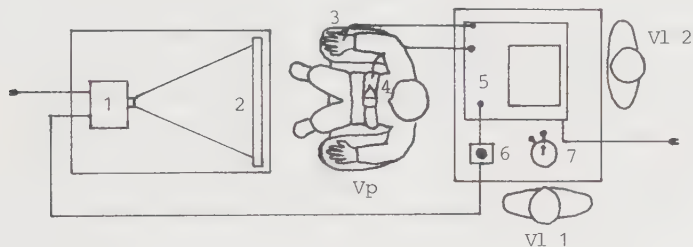


Abb. 17: Versuchsanordnung. 1 Diaprojektor, 2 Projektionsbildschirm, 3 Elektroden, 4 Atemgürtel, 5 Psychophysigraph, 6 Zweipoliger Taster, 7 Stoppuhr.

Bei der Versuchsbedingung mit zusätzlicher Kontrolle des Recency-Effekts wurde die Vp unmittelbar nach Ende der Diaprojektion aufgefordert, von der Zahl 925 ausgehend immer die Zahl 3 abzuziehen und das Ergebnis laut zu nennen. Nach 30 Sekunden wurde der Zählvorgang gestoppt. Anschließend wurden die Elektroden entfernt und die Reproduktionsprüfung und die weiteren Schritte wie schon beschrieben durchgeführt.

Sollte das Langzeitgedächtnis geprüft werden, wurde nach der Wortdarbietung mit der Vp ein Termin verabredet, an dem man ihr das Ergebnis der Kurvenauswertung mitteilen wollte. Beim zweiten Termin bekam die Vp im Anschluß an ihre Reproduktion zusätzlich die zugehörige Rekognitionsliste vorgelegt. Sie sollte die Wörter, an die sie sich noch erinnerte, ankreuzen.

Um die Anonymität der Vpn zu wahren, wurden nur ihr Vorname und ihre Vpn-Nummer aufgeschrieben, aber weder Nachname noch Adresse registriert. Dadurch konnte bei manchen Vpn die Langzeitgedächtnisprüfung nicht vorgenommen werden, da sie dem vereinbarten Termin fernblieben. Auf diese Weise fielen fünf Vpn aus.

4.3. Instruktionen

Unabsichtliches Lernen

In diesem Versuch sollen die physiologischen Reaktionen beim Lesen von Wörtern* untersucht werden. Auf dem Projektionsschirm werden Ihnen nun nacheinander etwas über 40 Wörter gezeigt. Sobald ein Wort erscheint, sollen Sie es still für sich lesen und auf sich wirken lassen. Konzentrieren Sie sich auf das Wort so lange es zu sehen ist. Aus meßtechnischen Gründen bleibt ein Wort etwa 5 sec sichtbar.

Absichtliches Lernen

Es handelt sich hier um ein Lernexperiment. In diesem Versuch sollen die physiologischen Reaktionen beim Lernen von Wörtern untersucht werden. Auf dem Projektionsschirm werden Ihnen nun nacheinander etwas über 40 Wörter gezeigt, die Sie sich einprägen sollen. Sobald ein Wort erscheint, sollen Sie es still für sich lesen und sich auf dieses Wort konzentrieren so lange es zu sehen ist. In diesem Experiment sollen die Wörter nicht mechanisch gelernt werden. Sie sollen also das Wort nicht ständig wiederholen, auch die Wörter nicht, die bereits gezeigt wurden, sondern immer nur das Wort auf sich wirken lassen, das gerade projiziert wird. Aus meßtechnischen Gründen bleibt ein Wort etwa 5 sec sichtbar. Im Anschluß an die Darbietung sollen Sie die Wörter, die Sie behalten haben, auf ein Blatt Papier schreiben.

* "und beim Zählen" (Zusatz bei der Bedingung mit der Zähl-
aufgabe; Kontrolle des Recency-Effekts)

4.4. Stichprobe

Vpn waren Studenten, einige Schüler und Berufstätige im Alter von 18 bis 45 Jahren (Durchschnittsalter 25 Jahre). Die Vpn wurden in der Universität angesprochen. Von 136 Vpn liegen Free Recall Daten vor (68 Männer, 68 Frauen). Die Stichprobe setzt sich folgendermaßen zusammen:

- 46% Psychologiestudenten
- 33% Studenten anderer Fächer
- 18% Berufstätige
- 3% Schüler.

Zahlenmäßige Verteilung der Vpn auf die einzelnen Versuchsgruppen:

	<u>männlich</u>	<u>weiblich</u>
absichtlich/Clusterliste	10	10
absichtlich/Willkürliste	10	10
ST unabsichtlich/Clusterliste	10	10
ST unabsichtlich/Willkürliste	10	10
ST Inc.Kontrolle/Willkürliste	11	12
LT unabsichtlich/Clusterliste	8	8
LT unabsichtlich/Willkürliste	9	8

5. Auswertung

5.1. Physiologische Daten (PHY)

Die vom Psychophysiographen aufgezeichneten Kurven wurden mit Hilfe einer selbstangefertigten Schablone ausgewertet (siehe Abb. 18). Die markierten Diawechsel wurden durchnummeriert (Wörter Nr. 1 - 46). Dann wurde die Schablone an die Markierung des jeweiligen Diawechsels angelegt und zu jeder Nummer notiert, ob während der Darbietung des Wortes eine elektrodermale Reaktion stattfand. Die Latenzzeit (LT, in sec) und die Höhe der Amplitude (in mm) wurde aufgeschrieben. Bei der letzten Markierung (Nr. 47), die den Lichtreiz anzeigte, konnte die Reaktionszeit der Vp festgestellt werden. In diesem Versuch lag die RT_{47} meistens bei 1 bis 2 sec (extreme Werte 0.9 bzw. 3.1 sec). Das Lesen eines Wortes beansprucht durchschnittlich 450 Millisekunden.

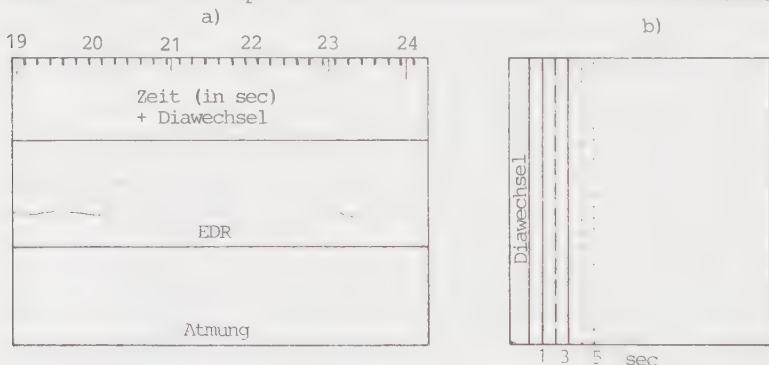


Abb. 18: a) Ausschnitt aus einer Kurvenaufzeichnung;
b) Schablone.

Nachdem jeder Kurvenverlauf transkribiert war, wurden nur jene SCRs gewertet, die eine LT von 1 bis 3 sec aufwiesen. Von diesen kann man annehmen, daß sie von den Wörtern ausgelöst worden sind.

Reaktionen, die während des Diawechsels oder bis 1/2 sec nach der Projektion auftraten, wurden nicht gewertet, ebenso

späte Reaktionen, d.h. solche mit einer Latenzzeit größer 4 1/2 sec. Die zugehörigen Wörter wurden alle als nicht erregend eingestuft.

Zu frühe Reaktionen mit einer LT von 0.5 bis 0.9 sec wurden als unsicher ausgeschieden. Auch Reaktionen mit einer LT größer 3 sec und kleiner 4.5 sec wurden als unsichere Fälle weggelassen, da hier nicht entscheidbar ist, ob die Bedeutung des Wortes erst spät erfaßt wurde und die SCR doch eine Reaktion auf das Wort war oder ob die Vp an andere Dinge dachte und so die Reaktion nichts mehr mit dem Testwort zu tun hatte. Ebenfalls nicht gewertet wurden Wörter, bei denen Meßstörungen vorlagen. Solche Störungen entstanden durch Bewegungen der Vp, durch Umschalten auf einen anderen Meßbereich oder wenn die Kurve außerhalb des Meßstreifens geraten war.

Wertungskriterien:

erregendes Wort (e): SCR mit LT 1 - 3 sec

neutrales Wort (n): keine SCR oder SCR mit LT <0.5 und >4.5

als unsicher ausgeschieden: SCR mit LT 0.5 - 0.9 und 3.1 - 4.4
oder Stellen mit Meßstörungen

Für jede Vp wurde eine Rangreihe der Wörter gebildet, Rang 1 erhielt das für die Vp am stärksten erregende Wort. An letzter Stelle wurden die neutralen Wörter aufgeschrieben. Die unsicheren Fälle wurden gesondert notiert, ebenso die Übungswörter.

Festgehalten wurde das Grundniveau der Hautleitfähigkeit, soweit ablesbar der Verlauf der Grunderregung ('Peaks'), der Verlauf der Atmung (regelmäßig oder unruhig) und ob viele, mittel viele oder wenige SCRs vorlagen. Weiterhin wurde registriert, bei welchen Nummern eine Veränderung der Atmung (tiefes Einatmen oder Atemanhalten) auftrat. Da tiefes Einatmen eine SCR auslöst, wurde die SCR nur berücksichtigt, wenn das Atmen auf das Wort erfolgte und die SCR innerhalb der Latenzzeit lag. Die Atmungsaufzeichnung diente als Kon-

trolle; sie ergab zusätzliche Hinweise auf die Erregung. Die Erregungsintensität der Wörter wurde nur an Hand der SCR bestimmt.

Die physiologischen Daten wurden nur ipsativ verarbeitet.

5.2. Subjektive Einschätzungen (SE)

Bei jeder Vp wurde für jedes Wort auf der Skala die Strecke zwischen Nullpunkt und Kennzeichnung in mm gemessen. Aus den ermittelten Werten wurde dann für jede Vp eine Rangreihe der Wörter gebildet (Rang 1 = stärkste Erregung).

Aufgrund der Einschätzungen der Vpn wurden für die Wörter der Clusterliste und der Willkürliste die mittleren empfundenen Erregungsstärken berechnet. Diese Daten dienten zur Überprüfung der a priori Kategorisierung.

5.3. Gedächtnisleistung

5.3.1. Free Recall Daten

Von jeder Vp wurde die Gesamtgedächtnisleistung ermittelt (Anzahl richtig reproduzierter Wörter der gesamten Liste/ 46 Wörter).

Es wurde festgestellt, wie oft jedes Wort von den Vpn reproduziert worden war. Daraus wurde eine Rangreihe erstellt; Rang 1 erhielt das am häufigsten erinnerte Wort.

Die Reihenfolge, in der die Wörter wiedergegeben wurden, sollte ebenfalls analysiert werden (entsprechend der Methode von MARX 1980). Hier war von Interesse, ob die E- und N-Wörter in getrennten Gruppen genannt wurden. Zusätzlich sollten die beiden Listenstrukturen überprüft werden: Wurden bei der Clusterliste die vorgegebenen Cluster gebildet? Erfolgte bei der Willkürliste keine Clusterbildung?

5.3.1.1. Auswertung nach dem Gruppenurteil

Vergleich E - N

Für jede Vp wurde bestimmt, wieviele E -und wieviele N - Wörter sie erinnert hatte. Da die Listen aus gleich vielen E- und N-Wörtern zusammengesetzt waren, konnte hier mit den absoluten Zahlen gerechnet werden. Die Daten wurden einer Varianzanalyse unterzogen.

5.3.1.2. Auswertung nach den physiologischen Daten

a) Vergleich e - n

Die erinnerten Wörter wurden im Auswertungsbogen der physiologischen Daten gekennzeichnet. Es wurde ausgezählt, wieviele Wörter für die jeweilige Vp erregend waren und wieviele sie von diesen erinnert hat; dann wieviele Wörter für die Vp neutral waren und wieviele sie hiervon erinnert hat. Manche Vpn wiesen mehr Wörter ohne Erregung auf, viele reagierten auf mehr als die Hälfte der Wörter mit einer Erregung. Der Fall, daß für eine Vp gleich viele physiologisch erregende wie physiologisch neutrale Wörter vorlagen, war äußerst selten. Bei der Zugrundelegung der physiologischen Daten konnte deshalb nur mit relativen Häufigkeiten gearbeitet werden.

Die Signifikanz des Unterschieds zwischen dem Anteil an e erinnerten und dem Anteil an n erinnerten Wörtern wurde mit der Formel für den kritischen Bruch (CR) von Prozentdifferenzen (HOFSTÄTTER & WENDT 1974, S.107) geprüft und mit dem Wilcoxontest. (Abhängige Stichproben)

Für den Vergleich der anderen Versuchsbedingungen (absichtlich - unabsichtlich; Clusterliste - Willkürliste; ST - LT) in ihrem Anteil an e erinnerten Wörtern wurde der U-Test von Mann-Whitney herangezogen. (Unabhängige Stichproben)

b) Vergleich stark e, mittel e, schwach e

Der Anteil erinnerter Wörter an den drei Bereichen wurde bestimmt. Ob ein Unterschied vorliegt, wurde mit Hilfe der Friedman-Rangvarianzanalyse geprüft. (Abhängige Stichproben)

5.3.1.3. Auswertung nach der subjektiven Einschätzung

Hier wurde in gleicher Weise vorgegangen wie bei der Auswertung nach den physiologischen Daten. Auch hier beurteilte die Mehrzahl der Vpn den größeren Teil der Wörter als erregend. Deshalb konnte nicht mit den absoluten Werten gerechnet werden.

5.3.2. Rekognitionsdaten

Zunächst wurde untersucht, welche Wörter wie häufig wiedererkannt wurden und welche Fehler begangen wurden (Items, die nicht in der Lernliste standen und trotzdem als wiedererkannt angestrichen wurden).

Weiter sollte geprüft werden, ob sich bei der Wiedererkennungslleistung der Erregungseffekt bemerkbar macht. Dabei wurden nur richtig gekennzeichnete Wörter verwendet.

6. Ergebnisse

6.1. Vorprüfungen

6.1.1. Gesamtgedächtnisleistung

Vorab wurde ermittelt, wie die Gesamtgedächtnisleistung bei den einzelnen Versuchsbedingungen ausfiel. Mit diesen Daten sollte geprüft werden, inwieweit die unabhängigen Variablen (Absichtlichkeit des Lernens, Zusammensetzung der Lernliste, Länge des Retentionsintervalls) wirksam waren.

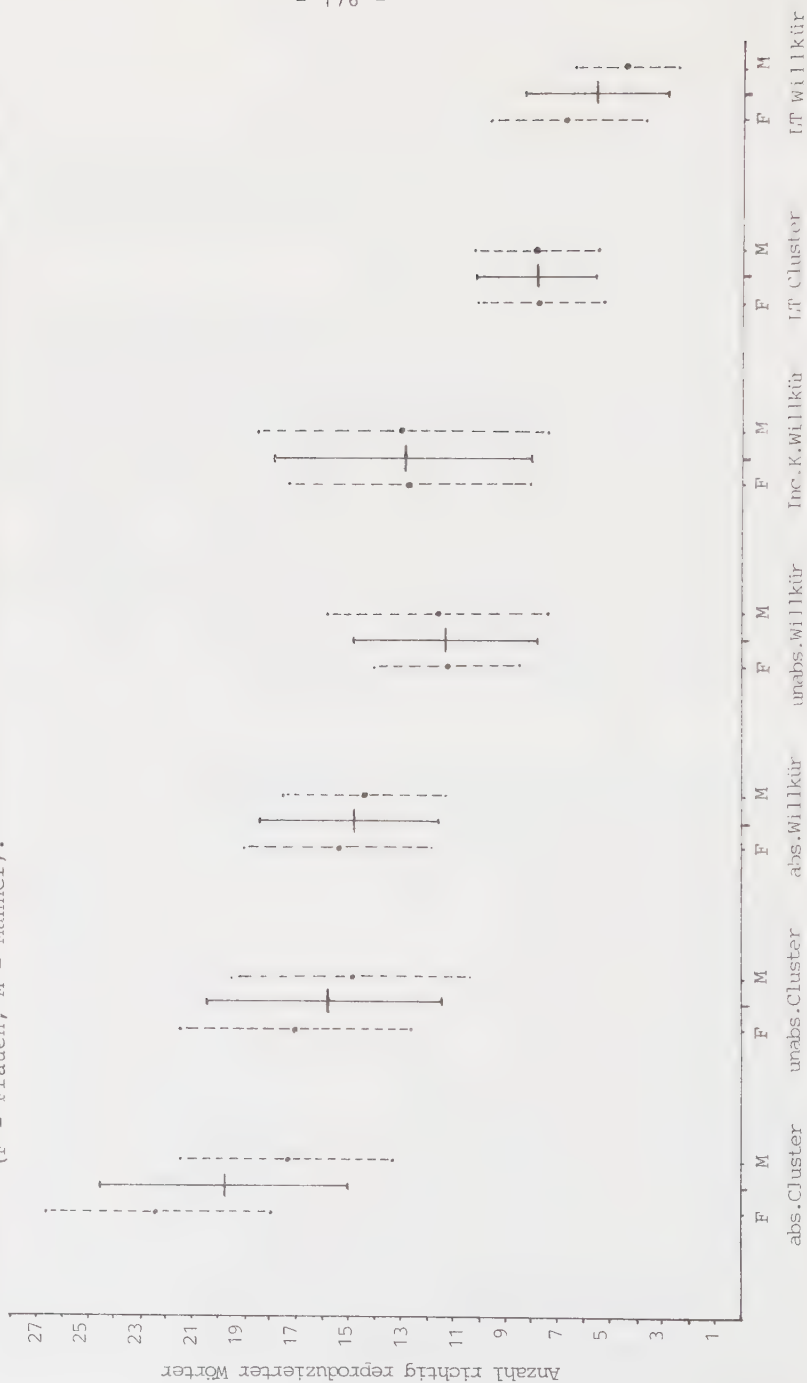
Erwartet wurde, daß die Vpn beim absichtlichen Lernen mehr Wörter reproduzieren als beim unabsichtlichen Lernen und daß sie von der Clusterliste mehr Wörter erinnern als von der Willkürliste. Ebenso sollte bei der Kurzzeitgedächtnisprüfung die Reproduktionsleistung größer ausfallen als bei der Langzeitgedächtnisprüfung.

In Tabelle 19 und in Abb. 19 befinden sich die Mittelwerte und Streuungen der einzelnen Versuchsgruppen.

	$\bar{X}$	σ		$\bar{X}$	σ
abs.Cluster (20 Vpn)	19.9	4.844	10 ♀	22.4	4.326
			10 ♂	17.4	4.12
unabs.Cluster (20 Vpn)	15.95	4.548	10 ♀	17	4.497
			10 ♂	14.9	4.58
abs.Willkür (20 Vpn)	14.9	3.386	10 ♀	15.4	3.66
			10 ♂	14.4	3.2
unabs.Willkür (20 Vpn)	11.3	3.555	10 ♀	11.1	2.77
			10 ♂	11.5	4.35
Inc.K.Willkür (23 Vpn)	12.78	4.963	12 ♀	12.67	4.58
			11 ♂	12.91	5.58
LT Cluster (16 Vpn)	7.8125	2.167	8 ♀	7.75	2.05
			8 ♂	7.875	2.42
LT Willkür (17 Vpn)	5.47	2.625	8 ♀	6.625	2.93
			9 ♂	4.44	1.94

Tabelle 19: Gesamtgedächtnisleistung der einzelnen Versuchsgruppen.

Abb. 19: Gesamtgedächtnisleistung (Mittelwert und Streuungen) der einzelnen Versuchsgruppen
(F = Frauen, M = Männer).



Die Mittelwertsunterschiede wurden mit dem t-Test für unabhängige Stichproben geprüft. Die erwarteten Unterschiede konnten bestätigt werden.

Unter der Bedingung 'absichtliches Lernen' werden signifikant mehr Wörter reproduziert als beim 'unabsichtlichen Lernen (einseitige Fragestellung).

absichtlich - unabsichtlich $t = 3.551$ (df = 78) $p < .0005$

Clusterliste: abs. - unabs. $t = 2.659$ (df = 38) $p < .01$

Willkürliste: abs. - unabs. $t = 3.279$ (df = 38) $p < .005$

Das absichtliche Lernen hat bei beiden Listen einen Einfluß auf die Behaltensleistung.

Von der Clusterliste werden signifikant mehr Wörter behalten als von der Willkürliste. Dies trifft sowohl für das absichtliche wie das unabsichtliche Lernen zu als auch für die Langzeitgedächtnisprüfung (einseitige Fragestellung).

Clusterliste - Willkürliste $t = 4.791$ (df = 78) $p < .0005$

abs.: Cluster - Willkürliste $t = 3.784$ (df = 38) $p < .0005$

unabs.: Cluster - Willkürliste $t = 3.602$ (df = 38) $p < .005$

LT: Cluster - Willkürliste $t = 2.786$ (df = 31) $p < .005$

Die Behaltensleistung fällt bei der unmittelbaren Behaltensprüfung wesentlich größer aus als bei der Langzeitgedächtnisprüfung (einseitige Fragestellung).

Clusterliste: ST - LT $t_{\text{welch}} = 7.063$ (f = 28) * $p < .0005$

Willkürliste: ST - LT $t = 5.583$ (df = 35) $p < .0005$

Beim unabsichtlichen Lernen der Willkürliste unterscheiden sich die Gesamtgedächtnisleistungen (Mittelwerte) nicht, gleichgültig ob der Recency-Effekt kontrolliert wird (Inc.K. Willkür) oder nicht (unabs. Willkür). Die Streuung ist bei der Störbedingung etwas größer (der F-Wert (1.95) ist aber nicht signifikant ($F_{\text{krit}} = 2.15$)). Einige Vpn erbrachten hier

* Zwischen ST und LT Cluster (unabs.Lernen) besteht keine Varianzhomogenität ($F = 4.4$; $F_{\text{krit}} = 2.36$), deshalb wurde hier das Verfahren von Welch zur Prüfung des Mittelwertunterschiedes bei ungleicher Varianz (CLAUSS/EBNER 1970, S.194) angewendet.

sehr schlechte Ergebnisse, fast auf dem Niveau der Langzeitgedächtnisprüfung. Für mehrere Vpn bedeutete das Rückwärtszählen eine starke Belastung. Dies zeigte auch das physiologische Maß der Hautleitfähigkeit deutlich. Das Grundniveau stieg an, und die phasischen Reaktionen erhöhten sich amplituden- und mengenmäßig.

Inc.K.Willkür - unabs. Willkür $t = 1.108$ ($df = 41$) n.s.
(zweiseitige Fragestellung)

Der Unterschied in der Gesamtgedächtnisleistung zwischen Frauen und Männern ist bis auf die Bedingung 'Clusterliste ST' nicht signifikant (zweiseitige Fragestellung).

Cluster ST: Frauen - Männer $t = 2.348$ ($df = 38$) $p < .05$

Bei der Clusterliste reproduzieren die Frauen bei der unmittelbaren Behaltensprüfung mehr Wörter als die Männer. Der Unterschied ist vor allem beim absichtlichen Lernen größer.

abs. Cluster: Frauen - Männer $t = 2.649$ ($df = 18$) $p < .02$

unabs.Cluster: Frauen - Männer $t = 1.035$ ($df = 18$) n.s.

6.1.2. Prüfung des Positionseffekts

Um zu prüfen, ob sich ein Positionseffekt bemerkbar gemacht hat, wurden die 46 Wörter der Lernliste wie folgt aufgeteilt:

I die sechs Übungswörter am Anfang der Liste,

dann die 40 Testwörter in vier Zehnergruppen:

II Wortposition 7 - 16

III Wortposition 17 - 26

IV Wortposition 27 - 36

V Wortposition 37 - 46

Die Anzahl reproduzierter Wörter (f_R) in den einzelnen Positionsgruppen und die Gesamtgedächtnisleistung ($\sum f_R$) wurde für jede Versuchsbedingung bestimmt. Aus der Gesamtgedächtnisleistung ($\sum f_R$) wurde der Erwartungswert für die Reproduktionshäufigkeit eines Items bei Annahme einer Gleichverteilung errechnet.

$$\text{Erwartungswert (Erw)} = \frac{\sum f_R}{46}$$

Für die sechs Übungswörter wurde der Erwartungswert mit 6 multipliziert (Erw_U), für die vier Zehnergruppen der Testwörter wurde der Erwartungswert mit 10 multipliziert (Erw_T).

Die beobachteten Werte (f_R) der einzelnen Positionsgruppen wurden dann durch die entsprechenden Erwartungswerte dividiert.

$$\frac{f_R}{Erw} = W$$

W gibt das Verhältnis von beobachtetem zu erwartetem Wert an.

$W = 1$: es besteht kein Unterschied zwischen beobachtetem und erwartetem Wert.

$W > 1$: es werden mehr Wörter erinnert als bei Annahme einer Gleichverteilung zu erwarten ist.

$W < 1$: es werden weniger Wörter erinnert als bei Annahme einer Gleichverteilung zu erwarten ist.

Fällt W deutlich größer oder kleiner als 1 aus, deutet das auf einen Positionseffekt hin.

In Tabelle 20 und 21 werden in der 1. Spalte die absoluten Werte der einzelnen Positionsgruppen (Anzahl reproduzierter Wörter, f_R) angegeben. (Die Werte der Übungs- und Testwörter sind nicht vergleichbar, da den Werten unterschiedliche Zahlen (6; 10) zugrundeliegen!) In der 2. Spalte stehen die Werte für W (f_R/Erw).

CLUSTERLISTE

	absichtlich (20 Vpn)		unabsichtlich (20 Vpn)		LT (16 Vpn)	
	f_R	W	f_R	W	f_R	W
I	32	.64	29	.70	13	.79
II	91	1.05	69	.99	35	1.29
III	90	1.04	64	.92	30	1.10
IV	99	1.14	74	1.07	30	1.10
V	85	.98	83	1.20	17	.63
	398	*	319		125	
Erw_U :	51.91		41.61		16.30	
Erw_T :	86.52		69.35		27.17	

* Zur Überprüfung eines Rechenfehlers sind die Werte von W der Positionen II - V zu addieren. Der Wert von W für die Übungsitens ist durch 10 zu dividieren und mit 6 zu multiplizieren. Zu der obigen Summe addiert, muß das Ergebnis 4.60 lauten.

CLUSTERLISTE

Frauen

	absichtlich (10 Vpn)		unabsichtlich (10 Vpn)		LT (8 Vpn)	
	f _R	W	f _R	W	f _R	W
I	22	.75	19	.86	7	.86
II	48	.99	40	1.08	16	1.19
III	50	1.03	35	.95	15	1.11
IV	54	1.11	38	1.03	14	1.04
V	50	1.03	38	1.03	10	.74
	224		170		62	
Erw _U :	29.22		22.17		8.09	
Erw _T :	48.70		36.96		13.48	

Männer

	absichtlich (10 Vpn)		unabsichtlich (10 Vpn)		LT (8 Vpn)	
	f _R	W	f _R	W	f _R	W
I	11	.48	10	.51	6	.73
II	43	1.14	29	.90	19	1.39
III	40	1.06	29	.90	15	1.09
IV	45	1.19	36	1.11	16	1.17
V	35	.99	45	1.39	7	.51
	174		149		63	
Erw _U :	22.70		19.43		8.22	
Erw _T :	37.83		32.39		13.70	

Tabelle 20: Häufigkeiten reproduzierter Items in den einzelnen Positionsgruppen der Clusterliste.

Abbildung 20 zeigt die relativen Häufigkeiten (W) der einzelnen Positionsgruppen bei der Clusterliste (abs., unabs., LT).

Die ersten 6 Übungswörter (I) wurden unter allen Versuchsbedingungen am seltensten erinnert, was sicher daran liegt, daß sie willkürlich ausgewählt waren, so daß sie weder mit den 10 Kategorien der Testliste noch untereinander einen kategorialen oder assoziativen Zusammenhang aufwiesen.

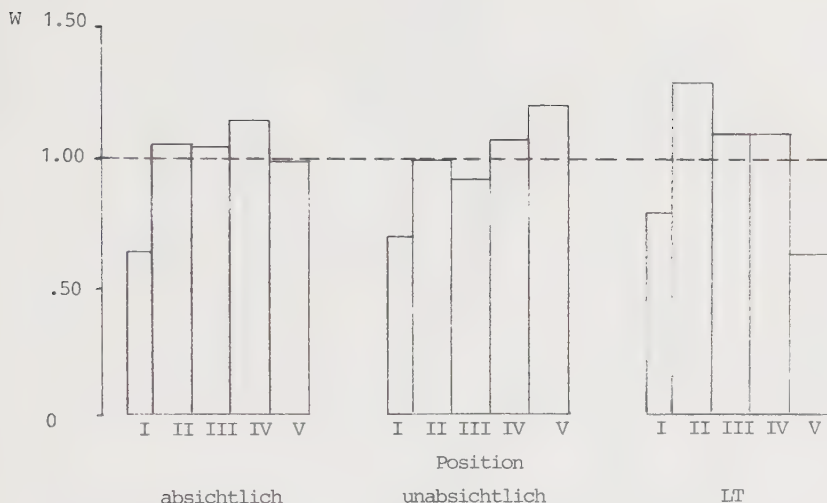


Abb. 20: Relative Häufigkeiten (W) in den einzelnen Positionen der Clusterliste.

Betrachtet man die Verteilung der Reproduktionen der Testwörter, ergeben sich folgende Unterschiede der drei Bedingungen.

Beim absichtlichen Lernen ist die Verteilung relativ gleichmäßig, ein Positionseffekt macht sich nicht bemerkbar.

Beim unabsichtlichen Lernen wurden die letzten 10 Items am häufigsten erinnert. Hier zeigt sich ein leichter Positionseffekt, der auf die Reproduktionsleistung der Männer zurückzuführen ist, während sich bei den Frauen die Reproduktionshäufigkeiten nicht unterscheiden. Vermutlich ist die Tendenz zum Clustern bei den Frauen deutlicher. Einen Hinweis liefert, neben der größeren Gesamtgedächtnisleistung bei der Clusterliste, der sehr deutliche Recency-Effekt beim unabsichtlichen Lernen der Willkürliste, der genauso stark ausgeprägt ist wie bei den Männern (Position V: Frauen 1.41, Männer 1.48, Tabelle 21).

Beim langfristigen Behalten schneiden die ersten Testwörter (Position II) besser ab, die Wörter am Ende der Liste (V) am schlechtesten.

WILLKÜRLISTE

Frauen

	absichtlich (10 Vpn)		unabsichtlich (10 Vpn)		LT (8 Vpn)	
	f _R	W	f _R	W	f _R	W
I	27	1.34	15	1.04	7	1.01
II	35	1.05	12	.50	8	.69
III	32	.96	25	1.04	12	1.04
IV	31	.93	25	1.04	15	1.30
V	29	.87	34	1.41	11	.95
	154		111		53	
Erw _Ü :	20.09		14.48		6.91	
Erw _T :	33.48		24.13		11.52	

Männer

	absichtlich (10 Vpn)		unabsichtlich (10 Vpn)		LT (9 Vpn)	
	f _R	W	f _R	W	f _R	W
I	24	1.28	20	1.33	8	1.53
II	22	.70	11	.44	9	1.03
III	36	1.15	19	.76	12	1.38
IV	34	1.09	28	1.12	7	.80
V	28	.89	37	1.48	4	.46
	144		115		40	
Erw _Ü :	18.78		15.00		5.22	
Erw _T :	31.30		25.00		8.70	

Inc.K.

	zusammen (23 Vpn)		Frauen (12 Vpn)		Männer (11 Vpn)	
	f _R	W	f _R	WW	f _R	W
I	57	1.49	27	1.36	30	1.62
II	48	.75	23	.70	25	.81
III	62	.97	31	.94	31	1.00
IV	66	1.03	40	1.21	26	.84
V	61	.95	31	.94	30	.97
	294		152		142	
Erw _Ü :	38.35		19.83		18.52	
Erw _T :	63.91		33.04		30.87	

	WILLKÜRLISTE					
	absichtlich (20 Vpn)		unabsichtlich (20 Vpn)		LT (17 Vpn)	
	f_R	W	f_R	W	f_R	W
I	51	1.31	35	1.19	15	1.24
II	57	.88	23	.47	17	.84
III	68	1.05	44	.90	24	1.19
IV	65	1.00	53	1.08	22	1.09
V	57	.88	71	1.46	15	.74
	298		226		93	
Erw _U :	38.87		29.48		12.13	
Erw _T :	64.78		49.13		20.22	

Tabelle 21 (Forts.): Häufigkeiten reproduzierter Items in den einzelnen Positionsgruppen der Willkürliste.

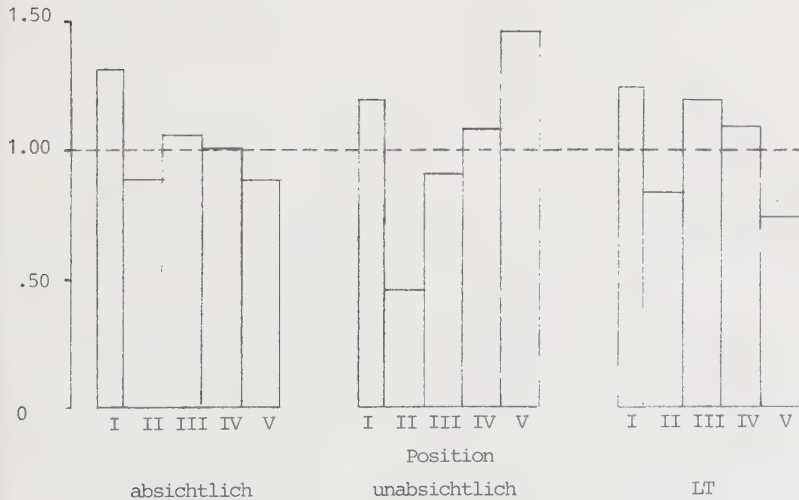


Abb. 21: Relative Häufigkeiten (W) in den einzelnen Positionsgruppen der Willkürliste.

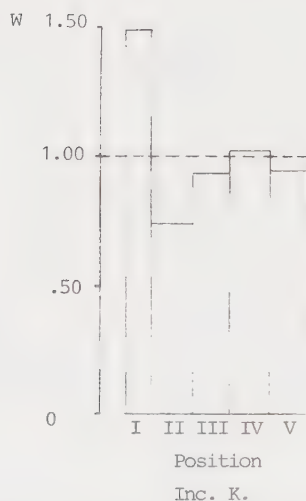


Abb. 21 (Forts.): Relative Häufigkeiten (W) in den einzelnen Positionsgruppen der Willkürliste.

Bei der Willkürliste werden unter allen Bedingungen die Übungswörter (I) vermehrt reproduziert. Dies ist nicht erstaunlich, da die Testwörter genauso willkürlich zusammengesetzt sind wie die Übungswörter, die aber durch ihre Anfangsposition hervorgehoben sind. Der Primacy-Effekt ist beim absichtlichen Lernen und bei der Bedingung 'Inc.K.' am größten. Beim unabsichtlichen Lernen und beim Langzeitgedächtnis tritt der Effekt nur bei den Männern und nicht bei den Frauen auf. Beim unabsichtlichen Lernen wurde dieser Effekt durch eine Vp verursacht, die sich die Übungswörter einprägte.

Die Verteilung der Testwörter ist auch hier beim absichtlichen Lernen ausgewogen.

Beim unabsichtlichen Lernen zeigt sich ein deutlicher Recency-Effekt. Die Wörter, die jeweils in der Schlußposition (V) stehen, werden am häufigsten reproduziert. Bei der Umkehrung der Reihenfolge befinden sich diese Wörter in Position II, wo sie am schlechtesten erinnert werden.

Wird aber beim unabsichtlichen Lernen im Anschluß an die

Darbietung eine kurze Störaufgabe eingeschoben, verschwindet der Recency-Effekt. Die letzten Wörter werden nicht mehr bevorzugt erinnert. Dafür werden aber vor allem die Übungswörter reproduziert. (Bedingung Inc.K.)

Beim Langzeitgedächtnis liegt eine ähnliche Verteilung vor wie beim absichtlichen Lernen. Die leichte Überlegenheit der Wörter in Position III und IV ist nicht auf ihre Lage zurückzuführen, sondern darauf, daß sich darunter die am stärksten erregenden Wörter befinden.

Bei der Prüfung des emotionalen Erregungspotentials (μV) werden die Übungswörter nicht berücksichtigt. Der Anfangspositionseffekt wird dadurch ausgeschlossen.

Da die Reihenfolge der Testwörter bei der Hälfte der Vpn umgekehrt wurde (Austausch Position II mit V und III mit IV) kann man die Positionseffekte über alle Vpn weitgehend vernachlässigen. Bei der ipsativen Betrachtung muß jedoch beachtet werden, daß innerhalb einer Vp Positionseffekte je nach Lernbedingung unterschiedlich stark zum tragen kommen. Vor allem gilt das für das Langzeitgedächtnis, für das unabsichtliche Lernen der Willkürliste und nur bei den Männern für das unabsichtliche Lernen der Clusterliste. Beim absichtlichen Lernen zeigten sich keine Positionseffekte bei den Testwörtern.

6.1.3 Überprüfung des Clusterings

Bei dieser Prüfung wird nach der "Methode zur geometrischen Darstellung des Clusterings bei Free-Recall-Daten" (MARX 1980) vorgegangen.

Die erinnerten Wörter einer Vp werden der reproduzierten Reihenfolge entsprechend von 1 bis n durchnummeriert (n = Gesamtzahl der reproduzierten Items incl. falsch reproduzierter Items). Dies wird mit allen Lernleistungen aller Vpn durchgeführt. Dann werden Distanzkoeffizienten (DK)

nach der Formel von MARX (S.166/167) berechnet.

$$DK_{ij} = \frac{\sum_{k=1}^{N_{ij}} E_{jk} - E_{ik}}{N_{ij}}$$

Dabei ist:

E_{jk} : Nummer des Elements j in der Reproduktionsreihe k

E_{ik} : Nummer des Elements i in der Reproduktionsreihe k

N_{ij} : Anzahl der Reproduktionsreihen, in denen die Elemente i und j gemeinsam vorkommen.

Leer gebliebene Zellen werden geschätzt durch den Mittelwert aller Distanzkoeffizienten der beiden Elemente.

Falsch reproduzierte Items und die Übungswörter wurden in diese Auswertung nicht miteinbezogen.

Im Anhang befinden sich die Matrizes der berechneten DK-Werte der beiden Lernlisten. Jede Matrix wurde einer non-metrischen multidimensionalen Skalierung (NMDS) unterzogen, die mit Hilfe des MINISSA-Programms von E.E.ROSKAM & J.C. LINGOES (Edinburgh Version) durchgeführt wurde. Es wurden für die Euklidische Metrik Lösungen mit 1, 2, 3, 4 und 5 Dimensionen gerechnet. Für die beiden Listen ergaben sich bei den einzelnen Dimensionslösungen folgende Stresswerte:

Dimensionen	Clusterliste	Willkürliste
	Stress $\hat{d}$	Stress $\hat{d}$
1	.46	.48
2	.28	.31
3	.19	.23
4	.14	.18
5	.12	.14

Alle Werte können als zufallskritisch abgesichert gelten (STENSON & KNOLL 1969, zit. nach MARX 1980).

Wegen der guten Anschaulichkeit fiel die Entscheidung jeweils für die zweidimensionale Lösung. Die Koordinaten der Elemente für die zweidimensionale Lösung sind im Anhang aufgeführt. Die geometrische Darstellung dieser Werte findet sich in Abb. 22 (Clusterliste) und Abb. 23 (Willkürliste).

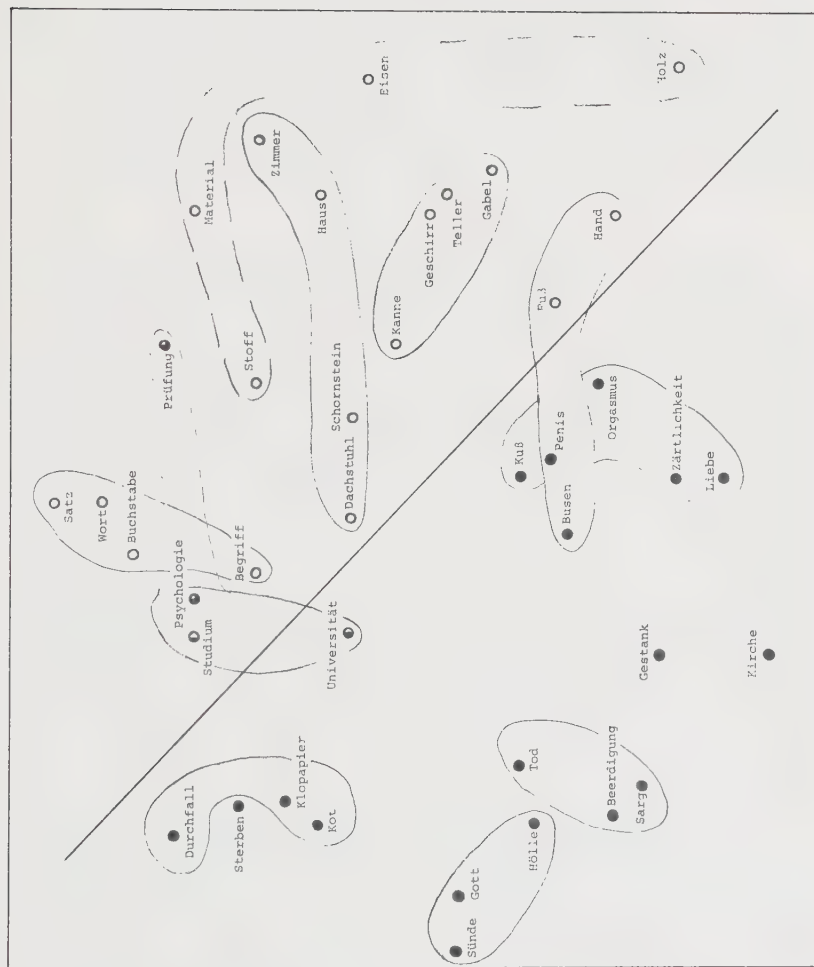


Abb. 22: Geometrische Darstellung der zweidimensionalen Euklidischen MINISSA-Lösung. (CLUSTERLISTE)

● emotionale Wörter ○ neutrale Wörter

Das vorliegende Ergebnis entspricht nicht in allen Punkten den intendierten Erwartungen. Zumindest drei der vier Wörter einer Kategorie werden jeweils enger beieinander reproduziert. Die Betrachtung von Abb. 22 macht einige wesentliche Aspekte deutlich:

Manche Kategorien weisen untereinander Verbindungen auf.

Daß die Kategorien 'Sprache' und 'Wissenschaft',

'Haushaltsgegenstände' und 'Haus',

'Körperteile' und 'Liebe',

'Religion' und 'Tod' näher beieinander sind, läßt sich aus ihrer assoziativen Zusammengehörigkeit (bezogen auf diese Liste) erklären.

Die Wörter der Kategorie 'Material' liegen weit auseinander, da die zugrundeliegenden Daten nicht sehr reliable sind.

Diese Wörter wurden äußerst selten reproduziert, deshalb mußten in den meisten Fällen Schätzwerte gebildet werden.

In höher dimensional Lösungen liegen diese Wörter näher beisammen.

Ausreißer:

Das Wort 'Prüfung' verschob sich hin zu 'Stoff' und 'Material'.

Eine Verbindung wird deutlich, wenn man an Wörter wie Lernstoff oder Lernmaterial denkt. Bei einigen Vpn, die kurz vor einer Prüfung standen, fiel auf, daß sie bei den Wörtern 'Prüfung', 'Material' und 'Stoff' einen deutlichen EDR-Ausschlag hatten.

Das Wort 'Sterben' taucht nahe beim Wort 'Klopapier' auf. Diese Verbindung wurde vermutlich durch das enge Aufeinanderfolgen der Wörter in der Lernliste (nur ein Wort dazwischen) verursacht.

Die Position des Wortes 'Gestank' ist nicht eindeutig zu klären, teilweise scheinen assoziative Verbindungen (Verweusung) eine Rolle zu spielen, andererseits könnte auch die Reihenfolge in der Lernliste verantwortlich sein.

Das Wort 'Kirche' tendiert zum Wort 'Liebe'. Es ist von den Kategorien 'Tod' und 'Liebe' etwa gleich weit entfernt. Entsprechende assoziative Bezüge lassen sich durchaus herstellen.

Eine Analyse der Verteilung der emotionalen und neutralen Wörter läßt eine deutliche Trennungslinie erkennen, die als Diagonale von links oben nach rechts unten führt. Rechts von dieser Linie befinden sich die neutralen Wörter, links die emotionalen. Bemerkenswert ist, daß durch diese Linie die neutralen von den emotional behafteten Körperteilen getrennt werden.

Willkürliste

Bei der Willkürliste ergeben sich keine Cluster. Die meisten Verbindungen lassen sich auf die Stellung der Wörter zueinander in der Lernliste zurückführen.

Interessant ist, daß die emotional behafteten Wörter diesmal mehr zur Mitte hin tendieren, während sich die neutralen Wörter am Rande befinden, annähernd kreisförmig um die emotionalen Wörter angeordnet.

Zusammenfassend kann man feststellen, daß die beiden Listenstrukturen bestätigt werden konnten und daß sich eine Trennung der emotionalen und neutralen Wörter in der Organisationsstruktur andeutet.

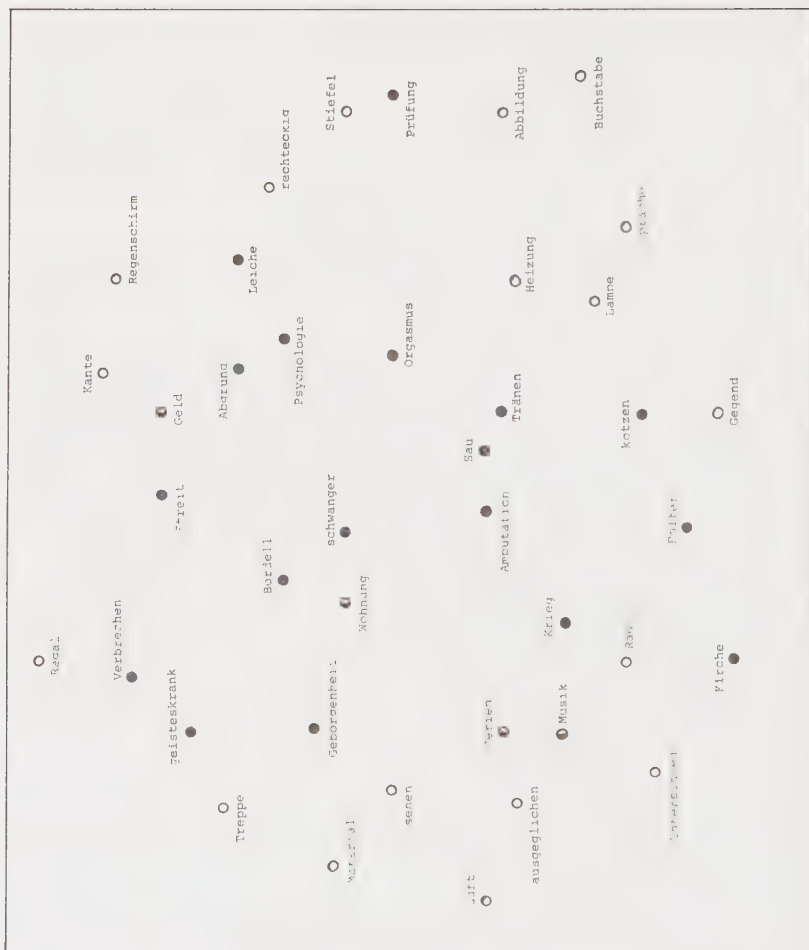


Abb. 23: Geometrische Darstellung der zweidimensionalen Euklidischen MINISSA-Lösung. (WILLKÜRLISTE)

● emotionale Wörter

○ neutrale Wörter

6.1.4. Subjektive Einschätzungen

Von den Testwörtern der beiden Lernlisten wurden durchschnittlich nur ca. 30% als eindeutig neutral eingestuft, der Rest erhielt mehr oder weniger starke Erregungsbeurteilungen.* Hier zeigt sich eine zahlenmäßige Entsprechung mit den EDR-Werten, ein Drittel der Testwörter rief keine EDR hervor.

Betrachtet man jede Liste für sich, lassen sich Unterschiede feststellen. Bei der Willkürliste werden durchschnittlich nur 10 Wörter als neutral beurteilt (Männer 8 Wörter; Frauen 11 Wörter). Für die Clusterliste ist der Durchschnittswert 15 Wörter (Männer 14 Wörter; Frauen 17 Wörter). Die männlichen Vpn sehen im Verhältnis zu den weiblichen Vpn etwas mehr Wörter als erregend an.

Zieht man jedoch die Höhe der beurteilten Erregungen heran, so wird deutlich, daß die vorgegebene Skala von den Frauen stärker in Richtung maximale Erregung genützt wird: Frauen vergeben höhere Werte für erregende Wörter als Männer. Der höchste Wert über alle weiblichen Vpn liegt bei 91.4 mm (Willkürliste) bzw. 88.1 mm (Clusterliste), die entsprechenden Höchstwerte der männlichen Vpn lauten 84.6 mm (Willkürliste) bzw. 71 mm (Clusterliste). Über alle Vpn sind die Werte 88.1 mm (Willkürliste) und 77.3 (Clusterliste).

Nur ein Drittel der Männer wählen bei der Willkürliste die maximale Erregung von 100 (mm), aber 44 % der Frauen. Auf die halbe Skala beschränkten sich lediglich 6 % der Frauen, aber 13 % der Männer. Bei der Clusterliste können für die Frauen keine wesentlichen Unterschiede festgestellt werden. Dagegen verwenden 40 % der Männer bei der Einschätzung der Wörter aus der Clusterliste die halbe Skala, und nur 17 % die ganze Skala.

* Bittet man die Vpn, die Wörter nur in die zwei Kategorien 'emotional' und 'neutral' zu sortieren, werden etwas mehr als die Hälfte der Wörter als neutral bezeichnet. Die schwach erregenden Wörter werden hier den neutralen zugeordnet. Der Vergleich der Rangreihen, die aus beiden Methoden entstehen, ergibt eine gute Übereinstimmung.

Die Wörter der Willkürliste werden insgesamt stärker erregend beurteilt als die Wörter der Clusterliste.

Die einzelnen Wörter werden teilweise sehr unterschiedlich bewertet. Was eine Vp als maximal erregendes Wort einschätzte, war für eine andere neutral. Beim Wort 'Tränen' reichen die Kennzeichnungen von 0 bis 100. Einige Wörter weisen nicht so große Schwankungen auf. Bis auf eine Ausnahme wurde das Wort 'Krieg' immer als mittel oder stark erregend klassifiziert, das Wort 'Kanne' wurde von allen Vpn als neutral oder sehr schwach erregend beurteilt.

In der folgenden Tabelle 22 sind die durchschnittlichen Beurteilungen (über alle Vpn) der einzelnen Wörter dargestellt, in einer Rangreihe geordnet, vom Wort mit dem höchsten bis zum Wort mit dem niedrigsten Erregungswert.

Willkürliste

Deutlich wird eine Übereinstimmung mit dem a priori Gruppenurteil. Alle neutralen Wörter befinden sich unter dem Median, alle emotionalen Wörter liegen darüber, abgesehen von den Wörtern 'ausgeglichen' (N) und 'Kirche' (E), die auf dem Median liegen. Das Wort 'ausgeglichen' scheint vermutlich für einen Teil der Vpn nicht so neutral zu sein, wie nach dem a priori Urteil angenommen werden konnte. Auch bei der physiologischen Messung rief es bei manchen Vpn Ausschläge hervor. Von der gemischten Gruppe (E/N) wurden zwei Wörter als stärker erregend beurteilt ('Musik' und 'Ferien'), an der großen Streuung lassen sich aber die unterschiedlichen Bewertungen ablesen; zwei Wörter ('Wohnung' und 'Geld') befinden sich knapp unter dem Median.

Vergleicht man die Daten von Männern und Frauen, zeigen sich bei 30 der 40 Wörter ähnliche Werte; drei Wörter wurden von den Frauen erregender beurteilt als von den Männern: 'Musik', 'schwanger' und 'Geborgenheit'. Wegen der großen Streuungen sind die Unterschiede aber nicht signifikant. Von den Männern wurden sieben Wörter erregender eingeschätzt: 'Gegend',

WILLKÜRLISTE	$\bar{x}$	σ^*	CLUSTERLISTE	$\bar{x}$	σ
1 Krieg (E)	67.4	25.8	1 Busen (E)	60.7	20.8
2 Orgasmus (E)	67.3	24.9	2 Kuß (E)	59.9	22.9
3 Musik (E/N)	63.9	30.1	3 Orgasmus (E)	59.3	27.2
4 Folter (E)	63.5	26.5	4 Zärtlichkeit (E)	57.7	25.9
5 Streit (E)	61.7	26.7	5 Liebe (E)	57.6	28.8
6 Amputation (E)	60.3	23.2	6 Penis (E)	55.0	24.3
7 Geborgenheit (E)	58.7	28.6	7 Tod (E)	50.1	22.7
8 Tränen (E)	58.2	29.2	8 Sterben (E)	49.0	23.7
9 Ferien (E/N)	58.1	33.1	9 Psychologie (E)	42.7	27.9
10 Leiche (E)	58.0	23.7	10 Prüfung (E)	38.3	23.5
11 geisteskrank (E)	53.5	26.2	11 Beerdigung (E)	37.1	22.1
12 schwanger (E)	53.1	25.7	12 Sarg (E)	30.2	21.7
13 Verbrechen (E)	51.7	29.6	13 Gott (E)	29.7	16.3
14 kotzen (E)	47.3	26.8	14 Gestank (E)	29.3	21.6
15 Abgrund (E)	46.6	28.9	15 Studium (N/E)	26.8	22.9
16 Prüfung (E)	45.4	26.6	16 Kot (E)	25.7	21.3
17 Psychologie (E)	44.5	23.4	17 Universität (E/N)	24.1	23.7
18 Bordell (E)	42.9	24.6	18 Durchfall (E)	21.7	20.7
19 Sau (E)	42.8	30.6	19 Kirche (E)	21.0	18.1
20 Kirche (E)	36.9	22.4	20 Hand (N)	20.9	24.9
21 ausgeglichen (N)	36.9	30.8	21 Hölle (E)	20.4	16.9
22 Wohnung (E/N)	35.5	28.2	22 Klopapier (E)	17.8	16.8
23 Geld (E/N)	27.8	24.2	23 Sünde (E)	16.0	18.9
24 Luft (N)	27.7	26.0	24 Holz (N)	10.8	19.5
25 Sprache (N)	26.3	27.8	25 Geschirr (N)	10.6	12.1
26 sehen (N)	23.4	24.5	26 Zimmer (N)	10.5	11.3
27 Gegend (N)	20.3	22.4	27 Schornstein (N)	9.8	18.6
28 Heizung (N)	20.3	23.1	28 Stoff (N)	9.5	19.4
29 Treppe (N)	16.6	19.5	29 Haus (N)	9.4	14.4
30 Kante (N)	16.5	25.7	30 Eisen (N)	8.6	17.3
31 Unterschied (N)	16.4	19.7	31 Fuß (N)	8.1	11.2
32 Rad (N)	15.4	21.6	32 Begriff (N)	7.1	11.6
33 rechteckig (N)	13.1	17.3	33 Wort (N)	6.3	9.4
34 Abbildung (N)	12.6	18.6	34 Buchstabe (N)	5.3	8.6
35 Material (N)	11.0	14.7	35 Satz (N)	5.0	6.2
36 Regal (N)	10.4	14.9	36 Gabel (N)	5.0	5.9
37 Stiefel (N)	9.7	12.8	37 Material (N)	4.9	7.1
38 Regenschirm (N)	9.1	11.7	38 Teller (N)	4.2	4.4
39 Buchstabe (N)	7.9	11.3	39 Dachstuhl (N)	4.1	6.0
40 Lampe (N)	7.4	10.5	40 Kanne (N)	3.8	3.7

Tabelle 22: Mittelwerte und Streuungen der subjektiven Einschätzungen auf der graphischen Ratingskala (Kontinuum von 100 mm Länge).

* Vermutlich hätte man die Streuungen durch Umwandlung der absoluten Werte der Vpn in relative etwas verringern können. Von meiner Seite bestand jedoch nur ein Interesse an den absoluten Werten.

'Kante', 'Regal', 'Sau', 'sehen', 'Unterschied' und 'Wohnung'. Die beiden Wörter 'Unterschied' und 'Regal' unterscheiden sich signifikant von den Werten der Frauen.

($t_{\text{Welch}} = 2.97$ bzw. 2.73 , $p < .05$)

Clusterliste

Insgesamt gesehen stimmen die Werte relativ gut mit dem a priori Urteil überein; 17 der 18 neutralen Wörter befinden sich unter dem Median (Ausnahme: 'Hand' knapp über dem Median). Von den E-Wörtern fallen diesmal drei knapp unter den Median ('Hölle', 'Klopapier', 'Sünde'), dafür rangieren die gemischt beurteilten Wörter über dem Median.

Die Kategorie 'Liebe' wurde am stärksten erregend eingestuft. Innerhalb der Kategorie 'Tod' wurden 'Tod' und 'Sterben' als erregender beurteilt als 'Beerdigung' und 'Sarg'. Die Kategorien 'Religion' und 'Toilette' nehmen einen mittleren Platz ein. Die gemischt zusammengesetzten Gruppen ('Körperteile' und 'Wissenschaft') zerfallen entsprechend der a priori Klassifikation in emotional behaftete und eher neutrale Wörter.

30 Wörter wurden von Frauen und Männern ähnlich beurteilt. Bei 10 Wörtern wählten die weiblichen Vpn höhere Erregungswerte: 'Durchfall', 'Hand', 'Kuß', 'Liebe', 'Penis', 'Psychologie', 'Sarg', 'Studium', 'Universität' und 'Zärtlichkeit'. Signifikant ist der Unterschied zu den Männern nur bei den Wörtern 'Hand' ($t_{\text{Welch}} = 2.69$, $p < .05$), 'Penis' ($t = 3.01$, $p < .01$) und 'Zärtlichkeit' ($t = 2.39$, $p < .05$).

Beide Listen hatten sechs Testwörter gemeinsam (Buchstabe, Kirche, Material, Orgasmus, Prüfung, Psychologie), die relativ gleich eingeschätzt wurden. Lediglich das Wort 'Kirche' erhielt in der Clusterliste signifikant niedrigere Intensitätswerte. ($t = 2.61$, $p < .02$)

6.2. Emotionales Erregungspotential und Free Recall

6.2.1. Emotionales Erregungspotential aufgrund des Gruppenurteils

Zur Überprüfung der Hypothese 1 wurden zwei Varianzanalysen (MANOVA) gerechnet. Die erste umfaßte die Variablen Absichtlichkeit des Lernens x Listenart x Geschlecht x Emotionales Erregungspotential, die zweite betraf die Variablen Länge des Retentionsintervalls x Listenart x Geschlecht x Emotionales Erregungspotential. Die Varianzanalyse prüfte je vier Faktoren, mit einer Meßwiederholung auf dem letzten Faktor (Programm SPSS/WINER 1971). Bei den zugrundeliegenden Daten handelt es sich nur um die Anzahl erinnerter E- und N-Wörter (ohne Übungswörter). Die Varianzanalyse wurde herangezogen, um zusätzlich zur Prüfung des emotionalen Erregungspotentials festzustellen, ob die verschiedenen anderen Bedingungen (absichtlich - unabsichtlich, Clusterliste - Willkürliste, ST - LT, Geschlecht) auch hinsichtlich der Behaltensleistung von E- und N-Wörtern unter Ausschluß der Übungswörter effektiv waren.

Varianzanalyse 1 (ST, abs. - unabs.)

Die Mittelwerte und Streuungen für die einzelnen Versuchsgruppen stehen in Tabelle 23.

	Vpn	E		N	
		$\bar{X}$	σ	$\bar{X}$	σ
absichtlich Cluster:	20	9.85	2.87	5.90	2.40
Frauen:	10	10.70	2.67	6.80	1.99
Männer:	10	9.00	2.94	5.00	2.54
unabsichtlich Cluster:	20	8.35	2.39	4.00	2.53
Frauen:	10	8.50	2.42	4.30	2.50
Männer:	10	8.20	2.49	3.70	2.67
absichtlich Willkür:	20	7.25	2.15	4.00	2.05
Frauen:	10	7.40	2.22	4.40	2.37
Männer:	10	7.10	2.18	3.60	1.71
unabsichtlich Willkür:	20	5.85	1.39	3.15	1.60
Frauen:	10	6.20	1.40	2.60	1.26
Männer:	10	5.50	1.35	3.70	1.77

	$\bar{X}_E$	$\bar{X}_N$
Clusterliste:	9.10	4.95
Willkürliste:	6.55	3.57
absichtliches Lernen:	8.55	4.95
unabsichtliches Lernen:	7.10	3.57
Frauen:	8.20	4.53
Männer:	7.45	4.00

Tabelle 23 (Forts.): Anzahl reproduzierter emotionaler (E) und neutraler Wörter (N).

Der Ergebnisse der Varianzanalyse 1 finden sich in Tabelle 24:

	SAQ	DF	MQ	F	Sign.
Error 1	441.95	72	5.72153		
Absichtlichkeit	79.80625	1	79.80625	13.94842	.00037
Listenart	154.05625	1	154.05625	26.92572	.00001
Geschlecht	16.25625	1	16.25625	2.84124	.09620
Abs. x Liste.	3.30625	1	3.30625	.57786	.44963
Abs. x Geschlecht	10.50625	1	10.50625	1.83627	.17963
Liste. x Geschl.	8.55625	1	8.55625	1.49545	.22536
Abs.x List.x Ges.	.75625	1	.75625	.13218	.71725
Error 2	292.75	72	4.06597		
Emotion.Erreg.	507.65625	1	507.65625	124.85482	.00001
Abs. x Emot.	.05625	1	.05625	.01383	.90670
Liste x Emot.	13.80625	1	13.80625	3.39556	.06949
Geschl. x Emot.	.50625	1	.50625	.12451	.72523
Abs.x List.x Emot.	2.25625	1	2.25625	.55491	.45874
Abs.x Ges. x Emot.	2.75625	1	2.75625	.67788	.41304
List.x Ges.x Emot.	1.85625	1	1.85625	.44414	.50722
Ab.x Li.x Ges.x Em.	3.90625	1	3.90625	.96072	.33029

Bis auf den Faktor 'Geschlecht' erweisen sich die Faktoren als sehr signifikant. Bei der Behaltensleistung von nur E- und N-Wörtern zeigt sich der gleiche Unterschied zwischen absichtlichem und unabsichtlichem Lernen wie bei der Analyse der Gesamtgedächtnisleistung, ebenso bei der Cluster- und Willkürliste. Unter den Bedingungen 'absichtliches Lernen'

und 'Clusterliste' werden mehr E+N Wörter reproduziert. Es liegen keine signifikanten Wechselwirkungen vor, lediglich bei den Faktoren Listenart x Emotionales Erregungspotential deutet sich eine Wechselwirkung an.

Die Variable 'Emotionales Erregungspotential' tritt hier deutlich in Erscheinung. Für die Kurzzeitgedächtnisleistung kann Hypothese 1 bestätigt werden. Es werden signifikant mehr emotionale Wörter (E) erinnert als neutrale (N), gleichgültig ob eine absichtliche oder unabsichtliche Lernbedingung vorliegt und ob es sich bei dem Lernmaterial um unverbundene oder um assoziativ verknüpfte Wörter handelt. Durchschnittlich wurden 7.83 emotionale und 4.26 neutrale Wörter wiedergegeben.

Varianzanalyse 2 (LT/ST, unabs.)

In Tabelle 25 sind die Mittelwerte und Streuungen der einzelnen Versuchsgruppen angegeben.

		E		N	
	Vpn	$\bar{X}$	σ	$\bar{X}$	σ
LT Clusterliste:	16	5.38	1.71	1.00	.89
Frauen:	8	5.50	1.93	.75	.71
Männer:	8	5.25	1.58	1.25	1.03
LT Willkürliste:	16	3.00	1.46	1.13	1.20
Frauen:	8	3.50	1.77	1.75	1.28
Männer:	8	2.50	.93	.50	.75
ST Clusterliste:	16*	7.94	2.43	3.88	2.12
Frauen:	8	8.25	2.66	4.00	2.00
Männer:	8	7.63	2.32	3.75	2.37
ST Willkürliste:	16	6.06	1.44	3.50	1.55
Frauen:	8	6.38	1.51	2.88	1.25
Männer:	8	5.75	1.39	4.13	1.64

* Bei der Langzeitgedächtnisbedingung waren nur Daten von je acht Vpn vorhanden. Deshalb wurden von der Bedingung Kurzzeitgedächtnis (unabsichtliches Lernen) jeweils die beiden zuletzt geprüften Vpn ausgeschieden, um gleich viele Vpn in jeder Zelle zu haben.

	$\bar{X}_E$	$\bar{X}_N$
Clusterliste:	6.66	2.44
Willkürliste:	4.53	2.31
LT:	4.19	1.06
ST:	7.00	3.69
Frauen:	5.91	2.34
Männer:	5.28	2.41

Tabelle 25 (Forts.): Anzahl reproduzierter emotionaler (E) und neutraler Wörter (N).

Tabelle 26 enthält die Ergebnisse der Varianzanalyse 2:

	SAQ	DF	MQ	F	Sign.
Error 1	186.375	56	3.32812		
LT/ST	236.53125	1	236.53125	71.07042	.00001
Listenart	40.50000	1	40.50000	12.16901	.00095
Geschlecht	2.53125	1	2.53125	.76056	.38688
LT/ST x Liste	.00000	1	.00000	.00000	1.00000
LT/ST x Gesch.	1.53125	1	1.53125	.46009	.50037
Liste x Gesch.	.50000	1	.50000	.15023	.69978
LT/ST x List.x Ge.	8.00000	1	8.00000	2.40376	.12668
Error 2	125.625	56	2.2433		
Emot.Erreg.	331.53125	1	331.53125	147.78706	.00001
LT/ST x Emot.	.28125	1	.28125	.12537	.72461
Liste x Emot.	32.00000	1	32.00000	14.26468	.00039
Ges. x Emot.	3.78000	1	3.78000	1.68557	.19951
LT/ST x List.x E.	2.00000	1	2.00000	.89154	.34912
LT/ST x Ges.x E.	1.53125	1	1.53125	.68259	.41220
List.x Ges. x E.	.12500	1	.12500	.05572	.81425
LT/ST x L.x G.x E.	3.12500	1	3.12500	1.39303	.24288

Die Variablen haben bis auf die Variable 'Geschlecht' eine sehr signifikante Wirkung. Bei der Langzeitgedächtnisprüfung werden weniger E+N Wörter reproduziert als bei der unmittelbaren Retentionsprüfung. Die Clusterliste schneidet wieder besser ab als die Willkürliste: die Behaltensleistung bzgl.

der E+N Wörter ist größer. Diesmal ist die Wechselwirkung zwischen Art der Liste und emotionalem Erregungspotential signifikant (siehe Abb. 24).

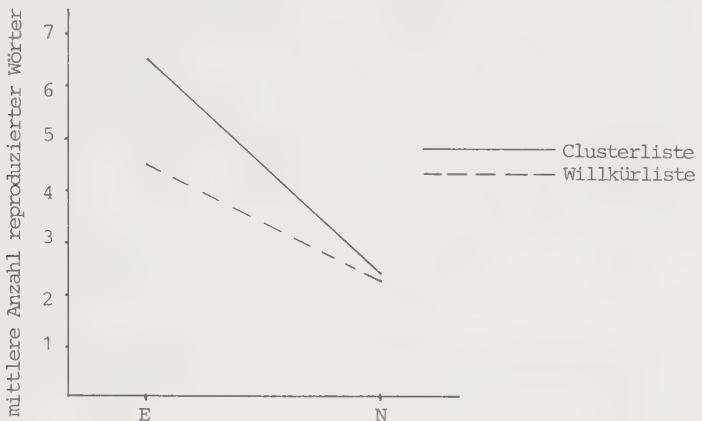


Abb. 24: Mittlere Reproduktionsleistung von E- und N- Wörtern der Cluster- und Willkürliste.

Die Behaltensmenge emotionaler Wörter ist bei der Clusterliste deutlich höher als bei der Willkürliste. Neutrale Wörter werden in beiden Listen mit gleicher Häufigkeit erinnert. Insgesamt werden mehr emotionale als neutrale Wörter wiedergegeben. Der Faktor 'Emotionales Erregungspotential' ist wiederum sehr signifikant. Es wurden durchschnittlich 5.6 emotionale und 2.38 neutrale Wörter reproduziert. Hypothese 1 trifft also auch für den Fall Langzeit- und Kurzzeitgedächtnis (unabsichtliches Lernen) zu.

Die Werte der Versuchbedingung 'Inc.Kontrolle Willk.' gingen in die Varianzanalyse nicht mit ein, weil bei keiner der anderen Versuchsbedingungen eine Störaufgabe gegeben wurde. Zur Prüfung der Hypothese 1 wurde hier der t-Test für abhängige Stichproben (einseitige Fragestellung) herangezogen. Tabelle 27 bringt neben den t-Werten für diese Bedingung zusätzlich die t-Werte der anderen Versuchsgruppen, die in diesem Zusammenhang ebenfalls berechnet wurden.

	t	df	p <
abs. Cluster:	5.52	19	.0005
Frauen:	4.46	9	.001
Männer:	3.38	9	.005
unabs. Cluster:	6.08	19	.0005
Frauen:	4.08	9	.005
Männer:	4.30	9	.001
abs. Willkür:	5.08	19	.0005
Frauen:	3.56	9	.005
Männer:	3.49	9	.005
unabs. Willkür:	6.91	19	.0005
Frauen:	6.91	9	.0005
Männer:	4.07	9	.005
Inc.K.Willk.:	5.05	22	.0005
Frauen:	2.22	11	.025
Männer:	6.77	10	.0005
LT Cluster:	9.08	15	.0005
Frauen:	6.54	7	.0005
Männer:	6.11	7	.0005
LT Willkür:	6.16	16	.0005
Frauen:	3.33	7	.01
Männer:	5.43	8	.0005

Tabelle 27: t-Werte der einzelnen Versuchsbedingungen (Prüfung des emotionalen Erregungspotentials).

Auch bei Einschieben einer Störaufgabe werden mehr emotionale als neutrale Wörter erinnert.

Betrachtet man die Werte von Männern und Frauen jeweils mit (Inc.K.Willk.) oder ohne (unabs.Willk.) Störbedingung, so wird eine Wechselwirkung sichtbar. Ohne Störaufgabe erinnern die Frauen deutlich mehr E-Wörter im Verhältnis zu den neutralen (6.2 vs. 2.6) als mit einer Störung (5.5 vs. 3.8). Bei den Männern ist es genau umgekehrt: sie reproduzieren unter der Störbedingung mehr E- und weniger N-Wörter (6.2 vs. 3.1; bei unabs.: 5.5 vs. 3.7).

Wird das emotionale Erregungspotential der Wörter anhand des Gruppenurteils bestimmt, so kann man Hypothese 1 als bestätigt ansehen: emotional erregende Wörter werden besser im

Gedächtnis behalten als neutrale Wörter, unabhängig davon welche Art der Lernliste vorgegeben wird, ob das Lernen absichtlich oder unabsichtlich erfolgt, wann die Behaltensprüfung stattfindet und ob eine Störbedingung vorliegt. Auch das Geschlecht spielt keine Rolle.

Im nächsten Schritt soll geprüft werden, ob sich der Anteil erregender Wörter unter den verschiedenen Lernbedingungen verändert.

Für jede Vp wurde der prozentuale Anteil der E-Wörter folgendermaßen berechnet:

$$\frac{E}{E + N} \cdot 100 *$$

Tabelle 28 gibt den durchschnittlichen prozentualen Anteil ($\bar{X}$) der E-Wörter unter den einzelnen Versuchsbedingungen zum Kurzzeitgedächtnis an.

	Vpn	$\bar{X}$		Vpn	$\bar{X}$
abs.Cluster	20	63.13%	Frauen	10	61.24%
			Männer	10	65.01%
unabs.Cluster	20	69.12%	Frauen	10	66.97%
			Männer	10	71.26%
abs.Willkür	20	65.39%	Frauen	10	64.35%
			Männer	10	66.43%
unabs.Willkür	20	66.26%	Frauen	10	71.36%
			Männer	10	61.16%

Tabelle 28: Prozentualer Anteil der emotionalen Wörter an der Gedächtnisleistung.

Die Daten wurden zwei Varianzanalysen (mit je drei Faktoren) unterzogen, in die erste gingen die Variablen Absichtlichkeit des Lernens, Listenart und Geschlecht ein, in die zweite Länge des Retentionsintervalls, Listenart und Geschlecht. Die Ergebnisse der ersten Varianzanalyse enthält Tabelle 29.

* Dieser Index wurde gewählt, weil bei einigen Vpn $N = 0$ war. Der Index E/N wäre hier ungünstig gewesen.

	SAQ	DF	MQ	F	Sign.
Error 1	1.36662	72	.01898		
Absichtlichkeit	.02369	1	.02369	1.24828	.26760
Listenart	.00016	1	.00016	.00843	.92709
Geschlecht	.00000	1	.00000	.00001	.99772
Abs. x Liste.	.01301	1	.01301	.68634	.41015
Abs. x Geschl.	.01743	1	.01743	.91820	.34116
Liste x Geschl.	.03266	1	.03266	1.72081	.19376
Abs.x List.x Ges.	.02037	1	.02037	1.07336	.30366

Tabelle 29: Ergebnisse der Varianzanalyse 1 über den Anteil emotionaler Wörter an der Gedächtnisleistung (ST).

Keiner der drei Faktoren übt eine signifikante Wirkung aus. Der Anteil erregender Wörter an der Behaltensleistung ist bei den verschiedenen Kurzzeitbedingungen relativ gleichbleibend. Es läßt sich eine schwache Tendenz der Daten in Richtung der formulierten Erwartung feststellen, wonach beim unabsichtlichen Lernen der Einfluß des emotionalen Erregungspotentials stärker in Erscheinung treten sollte. (Unabs.: 67.7%, abs.: 64.3%). Der Unterschied ist aber so klein, daß Hypothese 4 nicht bestätigt werden kann.

Hinsichtlich der Listenstruktur zeigen nur die Frauen einen leichten Anstieg entsprechend der Hypothese 5, daß bei der Willkürliste der Anteil an emotionalen Wörtern an der Reproduktion größer ist als bei der Clusterliste. Die Männer hingegen weisen einen höheren Wert bei der Clusterliste auf. Die Nullhypothese kann deshalb nicht zurückgewiesen werden.

Frauen und Männer unterscheiden sich bei der Kurzzeitgedächtnisleistung nicht in ihrem Anteil emotional erregender Wörter (66% bei beiden). Die Nullhypothese (Hypothese 8) muß beibehalten werden.

Die zweite Varianzanalyse (Ergebnisse siehe Tabelle 30) erbrachte eine signifikante Wirkung der Länge des Retentionsintervalls.

	SAQ	DF	MQ	F	Sign.
Error 1	1.18102	56	.02109		
LT/ST	.36589	1	.36589	17.34928	.00011
Listenart	.04802	1	.04802	2.27684	.13694
Geschlecht	.00011	1	.00011	.00527	.94239
LT/ST x Liste	.00556	1	.00556	.26358	.60969
LT/ST x Geschl.	.03666	1	.03666	1.73846	.19270
Liste x Geschl.	.01023	1	.01023	.48485	.48911
LT/ST x Li.x Ges.	.12009	1	.12009	5.69407	.02043

Tabelle 30: Ergebnisse der Varianzanalyse 2 über den Anteil emotionaler Wörter an der Gedächtnisleistung (LT + ST).

Die Werte (E%) für die einzelnen Versuchsgruppen befinden sich in Tabelle 31.

	Vpn	$\bar{X}$		Vpn	$\bar{X}$
LT Cluster	16	84.88%	Frauen	8	87.92%
			Männer	8	81.78%
LT Willkür	16	77.50%	Frauen	8	69.38%
			Männer	8	85.63%
ST Cluster	16	67.86%	Frauen	8	67.05%
			Männer	8	68.67%
ST Willkür	16	64.25	Frauen	8	69.58%
			Männer	8	58.92%

Tabelle 31: Prozentualer Anteil der emotionalen Wörter an der Gedächtnisleistung.

Beim Langzeitgedächtnis ist der Anteil der emotionalen Wörter bei der Reproduktion viel höher als bei der Kurzzeitgedächtnisprüfung (LT: 81%, ST: 66%). Hypothese 6 kann somit als bestätigt angesehen werden.

Die Variablen 'Listenstruktur' und 'Geschlecht' haben für sich genommen wiederum keinen Einfluß, aber diesmal kommt es zu einer Wechselwirkung aller drei Faktoren. (Abb. 25)

Von ST nach LT verzeichnen die Männer bei der Willkürliste eine größere Zunahme im Anteil an E als bei der Clusterliste, im Gegensatz zu den Frauen, die bei der Willkürliste von ST

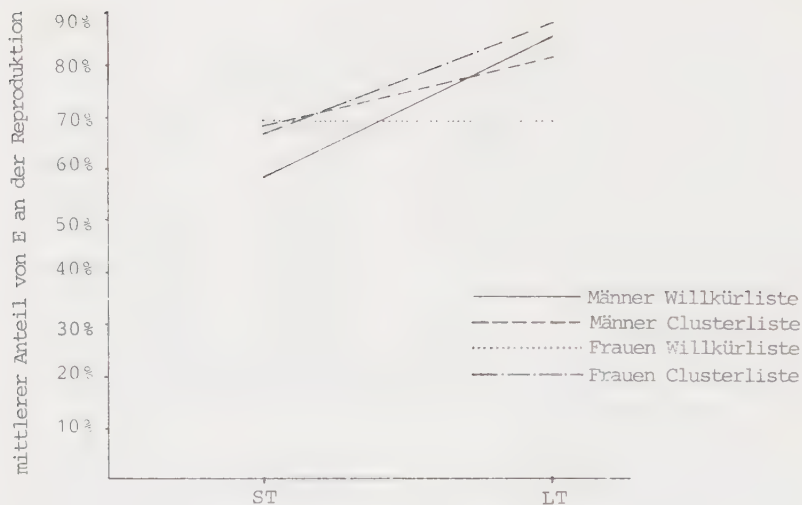


Abb. 25: Mittlerer Anteil von E (%) an der Reproduktion.

nach LT gleichbleibende Werte* haben und die nur einen Zuwachs bei der Clusterliste aufweisen.

Zur Prüfung von Hypothese 7 wurde der U-Test von Mann-Whitney herangezogen. Der Vergleich der Bedingungen 'unabs. Willkür' und 'Inc.K. Willkür' ergab keinen signifikanten Unterschied ($z = .17$, n.s.). Der Anteil emotionaler Wörter war mit Störung (66.40%) etwa genauso hoch wie ohne Störung (66.26%). Bei einer Differenzierung nach Geschlecht zeigt sich dieselbe Wechselwirkung zwischen Geschlecht und Störaufgabe wie bei den absoluten Werten (S.200 dieser Arbeit). Bei den Frauen ist der Anteil erregender Wörter an der Reproduktion ohne Störung größer (71.36%) als mit Störung (62.15%). Dieser Un-

* Bei der Bedingung 'längerfristiges Behalten' (Cluster- und Willkürliste) ergab sich in vier Fällen ein Störfaktor. Die betreffenden Vpn stellten nach Ende der Wortdarbietung Fragen zu bestimmten (meist neutralen) Wörtern. Diese Wörter wurden vermutlich durch die Wiederholung besser behalten, denn diese Wörter wurden unter anderen bei der Retentionsprüfung reproduziert.

terschied erweist sich aber nicht als signifikant ($U = 37.5$, $U_{krit.} = 29$, $p = .05$, zweiseitige Fragestellung). Männer verhalten sich hier entsprechend der formulierten Erwartung, ihr E-Anteil fällt mit Störung größer aus (71.03%) als ohne Störung (61.16%). Diese Tendenz ist knapp nicht signifikant ($U = 28.5$, $U_{krit.} = 26$). Hypothese 7 kann somit nicht bestätigt werden.

Am Ende dieses Abschnitts sollen die Wörter der Clusterliste und der Willkürliste in der Reihenfolge ihrer Reproduktionshäufigkeit angeführt werden (Tabelle 32).

Die Daten basieren auf je 56 Vpn. Bei der Willkürliste werden zusätzlich in Klammern die Werte bei Zugrundelegung von 80 Vpn (Daten incl. Inc.K. + 1 Vp LT) angegeben.

In beiden Listen befindet sich der Großteil der E-Wörter über dem Median. Von den N-Wörtern liegen aus der Willkürliste 'Kante*', 'Lampe*', 'ausgeglichen', 'Treppe' über dem Median, aus der Clusterliste die Wörter 'Haus', 'Wort', 'Fuß'. Schlechter erinnert wurden von der Willkürliste die E-Wörter 'Streit', 'Tränen', 'geisteskrank' und die gemischt beurteilten Wörter 'Ferien', 'Geld', von der Clusterliste die E-Wörter 'Durchfall', 'Hölle', 'Sünde', 'Klopapier', 'Gestank'.

Auffällig war, daß von der Willkürliste die weiblichen Vpn häufiger die Wörter 'Orgasmus', 'schwanger', 'Streit', 'ausgeglichen' reproduzierten, die männlichen Vpn dagegen die Wörter 'Psychologie', 'Bordell', 'Abgrund', 'Kirche'.

Bei der Clusterliste erinnerten die Frauen öfter die Wörter 'Prüfung', 'Sterben', 'Zimmer', die Männer das Wort 'Kuß'.

Die sechs Testwörter (Buchstabe, Kirche, Material, Orgasmus, Prüfung, Psychologie) unterscheiden sich kaum in ihrer Reproduktionshäufigkeit. Die Listenstruktur scheint die Reproduktionswahrscheinlichkeit dieser Wörter nicht zu beeinflussen.

* Diese Wörter wurden fast ausnahmslos erinnert, wenn sie in der Schlußposition standen.

WILLKÜRLISTE

- 44 Orgasmus (63) (E)
- 32 Psychologie (41) (E)
- 28 Bordell (39) (E)
- 23 Krieg (33) (E)
- 20 Sau (29) (E)
- 17 schwanger (22) (E)
- 16 Prüfung (25) (E)
- Geborgenheit (20) (E)
- 15 Kirche (20) (E)
- 14 Kante (20) (N)
- 13 Leiche (20) (E)
- kotzen (19) (E)
- Amputation (18) (E)
- Lampe (16) (N)
- 12 Abgrund (16) (E)
- Musik (16) (E/N)
- ausgeglichen (18) (N)
- Treppe (17) (N)
- 11 Folter (16) (E)
- Verbrechen (18) (E)
- Wohnung (18) (E/N)
- Regal (13) (N)
- rechteckig (17) (N)
- 10 Abbildung (14) (N)
- 9 Streit (20) (E)
- Ferien (12) (E/N)
- Rad (11) (N)
- Material (10) (N)
- 8 Tränen (9) (E)
- Regenschirm (16) (N)
- sehen (12) (N)
- 7 Gegend (13) (N)
- Sprache (9) (N)
- 6 Stiefel (7) (N)
- Unterschied (7) (N)
- 5 Heizung (13) (N)
- Buchstabe (11) (N)
- Luft (8) (N)
- 4 geisteskrank (5) (E)
- 3 Geld (6) (E/N)

CLUSTERLISTE

- 48 Penis (E)
- 43 Orgasmus (E)
- 39 Busen (E)
- Tod (E)
- 35 Beerdigung (E)
- 34 Zärtlichkeit (E)
- 31 Universität (E/N)
- 28 Liebe (E)
- 27 Psychologie (E)
- Studium (N/E)
- 26 Kuß (E)
- Sarg (E)
- 25 Haus (N)
- 24 Gott (E)
- 19 Kot (E)
- 18 Kirche (E)
- Prüfung (E)
- Fuß (N)
- Wort (N)
- 17 Sterben (E)
- Gabel (N)
- 14 Geschirr (N)
- Teller (N)
- 13 Durchfall (E)
- 12 Hölle (E)
- Eisen (N)
- 11 Sünde (E)
- Begriff (N)
- Dachstuhl (N)
- Kanne (N)
- 10 Schornstein (N)
- Stoff (N)
- Zimmer (N)
- 9 Klopapier (E)
- Hand (N)
- Satz (N)
- 7 Gestank (E)
- Material (N)
- 4 Buchstabe (N)
- Holz (N)

498 $\bar{X}$ 12.45

(714) ($\bar{X}$ = 17.85)

765 $\bar{X}$ = 19.125

Tabelle 32: Reproduktionshäufigkeit der Testwörter.

6.2.2. Emotionales Erregungspotential aufgrund der physiologischen Daten

Von den 136 Kurven waren 129 verwendbar. Bei sieben Vpn konnte die Kurve aus folgenden Gründen nicht ausgewertet werden: zweimal lösten sich während des Versuchs die Elektroden, dreimal lagen Störungen von seiten des Meßgeräts vor und zwei weibliche Vpn produzierten nach den Übungswörtern keine Ausschläge mehr.

Durchschnittlich waren 33 der 40 Testwörter eindeutig auswertbar (82%). Der Rest fiel weg entweder aufgrund von Meßstörungen oder wegen nicht eindeutiger Zuordnung.

Bei den meisten Vpn riefen mehr als die Hälfte der Wörter Ausschläge hervor. 17 von 64 weiblichen Vpn wiesen mehr neutrale Wörter (Wörter ohne EDR) auf, bei den Männern dagegen nur 5 von 65 Vpn.

In Tabelle 33 finden sich die mittlere Anzahl der Wörter mit (e) und ohne (n) EDR-Ausschlag und die dazugehörige Anzahl von Reproduktionen (Mittelwert und Prozent*).

	Vpn	e			n		
		$\bar{X}$	$\bar{X}$	%	$\bar{X}$	$\bar{X}$	%
abs.Cluster:	18	21.39	11.89	55.58%	9.06	2.89	31.90%
Frauen:	9	20.22	12.33	60.99%	11.78	4.44	37.74%
Männer:	9	22.55	11.44	50.74%	6.33	1.33	21.05%
unabs.Cluster:	20	22.65	9.90	43.71%	11.10	2.25	20.27%
Frauen:	10	21.20	9.90	46.92%	11.70	2.20	18.80%
Männer:	10	24.20	9.90	40.91%	10.50	2.30	21.90%
abs.Willkür:	17	22.06	8.59	38.93%	8.59	1.12	13.01%
Frauen:	9	20.67	8.89	43.01%	10.56	1.11	10.53%
Männer:	8	23.63	8.25	34.92%	6.38	1.13	17.65%
unabs.Willkür:	19	21.68	5.89	27.18%	11.79	2.16	18.30%
Frauen:	9	18.77	5.44	28.99%	16.44	2.89	17.57%
Männer:	10	24.30	6.30	25.93%	7.60	1.50	19.74%

* $\frac{\text{Anzahl erinnerter Wörter der jeweiligen Kategorie}}{\text{Gesamtzahl der Wörter dieser Kategorie}} \cdot 100$

	Vpn	e	erinnert		n	erinnert	
		$\bar{X}$	$\bar{X}$	%	$\bar{X}$	$\bar{X}$	%
Inc.K.Willkür:	23	23.96	7.96	33.21%	10.22	1.13	11.06%
Frauen:	12	24.00	8.16	34.03%	9.42	.92	9.73%
Männer:	11	23.91	7.70	32.32%	11.09	1.36	12.30%
LT Cluster:	16	22.06	5.81	26.35%	9.31	.44	4.70%
Frauen:	8	18.50	5.38	29.35%	13.50	.75	5.56%
Männer:	8	25.63	6.25	24.39%	5.13	.13	2.44%
LT Willkür:	16	19.56	3.50	17.89%	14.94	.56	3.77%
Frauen:	7	19.85	4.29	21.58%	15.00	.86	5.71%
Männer:	9	19.33	2.88	14.94%	14.89	.33	2.24%

Tabelle 33 (Forts.): Mittlere Anzahl von e und n Wörtern und mittlere Anzahl erinnerter Wörter in jeder Kategorie.

Hypothese 1

Die Prozentwerte dieser Tabelle für e und n wurden mit der Formel für den Kritischen Bruch (CR) von Prozent-Differenzen geprüft (HOFSTÄTTER & WENDT 1974, S.107). (Ergebnisse siehe Tabelle 34)

	CR	df	p
abs.Cluster:	5.330	546	.0005
Frauen:	3.917	286	.0005
Männer:	4.611	258	.0005
unabs.Cluster:	6.575	673	.0005
Frauen:	5.641	326	.0005
Männer:	3.708	345	.0005
abs.Willkür:	6.905	519	.0005
Frauen:	6.759	279	.0005
Männer:	2.713	238	.005
unabs.Willkür:	2.621	634	.005
Frauen:	2.437	315	.01
Männer:	1.154	317	n.s.
Inc.K.Willkür:	7.730	784	.0005
Frauen:	6.159	399	.0005
Männer:	4.835	383	.0005
LT Cluster:	7.424	500	.0005
Frauen:	5.419	254	.0005
Männer:	5.705	244	.0005
LT Willkür:	5.666	550	.0005
Frauen:	3.815	242	.0005
Männer:	4.248	306	.0005

Tabelle 34: CR-Werte der einzelnen Versuchsgruppen

Für alle sieben Vpn-Gruppen sind die Unterschiede im Anteil an 'e erinnert' und 'n erinnert' sehr signifikant. Es wurden mehr von den physiologisch erregenden Wörtern behalten als von den physiologisch nicht-erregenden Wörtern. Auch bei der getrennten Betrachtung der Werte von Frauen und Männern innerhalb jeder Gruppe zeigen sich bis auf eine Ausnahme ebenfalls hoch signifikante Unterschiede. Nur in der Gruppe 'unab.Willkür' (männliche Vpn) ist die Differenz nicht signifikant. Dies ist auf den Positionseffekt zurückzuführen, der hier stark zum tragen kam (vor allem die letzten Wörter wurden erinnert). Bei der Gruppe 'Inc.K.Willkür', wo der Positionseffekt durch die Störaufgabe ausgeschaltet wurde, ergibt sich ein hoch signifikanter Effekt zwischen 'e erinnert' und 'n erinnert'.

Die Verwendung des kritischen Bruchs (CR) ist in diesem Fall kein ganz unbedenkliches Verfahren (Fehlen einer echten Normalverteilung?), deshalb wurden die Daten auch noch mit Hilfe eines nonparametrischen Verfahrens (Wilcoxonstest) unter Zugrundelegung der individuellen Prozentwerte analysiert. Wieder sind die Ergebnisse der Gesamtgruppen hoch signifikant ($p < .005$; nur bei 'unabs. Willkür' $p < .025$). Bei drei Untergruppen finden sich diesmal keine signifikanten Unterschiede: 'unabs.Willkür Männer' ($T = 12$, $N = 10$ n.s.) (wie zuvor), zusätzlich 'unabs.Willkür Frauen' ($T = 8$, $N = 9$ n.s.) und 'abs.Willkür Männer' ($T = 5.5$, $N = 8$ n.s.).

Basiert die Bestimmung der Emotionalität der Wörter auf physiologischen Daten (EDR), so kann Hypothese 1 mit einer Ausnahme ('unabs.Willkür Männer' und 'unabs.Willkür Frauen') als bestätigt angesehen werden. Bei Ausschaltung des Recency-Effekts trifft Hypothese 1 ausnahmslos zu, es werden mehr emotionale (e) als neutrale (n) Wörter erinnert.

Um zu ermitteln, ob sich der Anteil der 'e erinnerten' Wörter an der Gedächtnisleistung unter den verschiedenen Lernbedingungen verändert, war eine weitere Umwandlung der Daten

erforderlich. Da die Gedächtnisleistung (e erinnert + n erinnert) für die einzelnen Gruppen (und Vpn) unterschiedlich ausfällt, ist es für einen Vergleich über die Gruppen (und Vpn) nötig, die absoluten Werte in relative umzuwandeln. Es wurde deshalb der prozentuale Anteil der 'e erinnerten' Wörter an der Gedächtnisleistung (e erinnert + n erinnert) berechnet.

prozentualer Anteil von 'e erinnert' an der Gedächtnisleistung =
$$\frac{e \text{ erinnert } \%}{e \text{ erinnert } \% + n \text{ erinnert } \%} \cdot 100$$

Tabelle 35 gibt den prozentualen Anteil von 'e erinnert' an der Gedächtnisleistung wieder.

	Vpn	Anteil e		Vpn	Anteil e
abs.Cluster:	18	63.53%	Frauen:	9	61.77%
			Männer:	9	70.68%
unabs.Cluster:	20	68.32%	Frauen:	10	71.39%
			Männer:	10	65.13%
abs.Willkür:	17	74.95%	Frauen:	9	80.33%
			Männer:	8	66.43%
unabs.Willkür:	19	60.35%	Frauen:	9	62.26%
			Männer:	10	56.78%
Inc.K.Willkür:	23	75.02%	Frauen:	12	77.77%
			Männer:	11	72.44%
LT Cluster:	16	84.86%	Frauen:	8	83.94%
			Männer:	8	90.91%
LT Willkür:	16	82.59%	Frauen:	7	79.08%
			Männer:	9	86.96%

Tabelle 35: Prozentualer Anteil 'e erinnerte' Wörter an der Gedächtnisleistung.

Der Vergleich der verschiedenen Lernbedingungen in ihrem Anteil an 'e erinnert' wurde mit dem U-Test von Mann und Whitney vorgenommen.

Hypothese 4 (Vergleich abs. - unabs.)

Es zeigt sich kein signifikanter Unterschied im Anteil an 'e erinnert' zwischen absichtlichem und unabsichtlichem Lernen, einzige Ausnahme: 'Willkürliste ohne Störaufgabe'. Hier

fallen die Prozentwerte entgegen der Annahme bei der Bedingung 'absichtliches Lernen' signifikant höher aus ($U = 87$, $U_{krit.} = 88$ $p = .02$, zweiseitige Fragestellung). Verursacht wird dieses Ergebnis durch den Recency-Effekt. Hypothese 4 trifft also nicht zu. Läßt man das als Artefakt einzustufende Ergebnis (unabs. Willkür) außer acht, so kann festgestellt werden, daß ebenso wie bei der Analyse des emotionalen Erregungspotentials nach dem Gruppenurteil kein Unterschied zwischen absichtlichem und unabsichtlichem Lernen vorliegt.

Hypothese 5 (Vergleich Clusterliste - Willkürliste)

Clusterliste und Willkürliste unterscheiden sich nicht bei der Kurzzeit- und Langzeitbehaltensprüfung bzgl. ihres Anteils an 'e erinnert'. Nur beim absichtlichen Lernen schneidet die Willkürliste besser ab ($U = 93$, $U_{krit.} = 93$ $p = .05$). Dies liegt an den Werten der Frauen, die bei der Clusterliste einen niedrigeren 'e' Anteil aufweisen. ($U = 10$, $U_{krit.} = 14$ $p = .02$). Vermutlich nützten die Frauen die Struktur der Liste besser aus und orientierten sich mehr an Zusammenhängen, wodurch der Faktor des emotionalen Erregungspotentials weniger wirksam wurde. Hypothese 5 kann nur für die Bedingung 'absichtliches Lernen, Frauen' als bestätigt angesehen werden. Für die anderen Lernbedingungen muß die Nullhypothese beibehalten werden.

Hypothese 6 (Vergleich LT - ST)

Ein sehr signifikanter Unterschied ($p < .00003$ $z = 4.19^*$) besteht im Anteil an 'e erinnert' zwischen Langzeitgedächtnis und Kurzzeitgedächtnis. Wie erwartet ist der Anteil beim Langzeitgedächtnis größer. Bei der Kurzzeitgedächtnisprüfung wird die Wirkung des emotionalen Erregungspotentials noch nicht so deutlich, da kurzfristig auch neutrales Material präsent ist, das bei der Langzeitgedächtnisprüfung nicht mehr reproduzierbar ist. Die Stimuli, die den Organismus

* z wurde nach Formel 6.9 (SIEGEL 1976, S.120) berechnet, die Verbundwerte berücksichtigt.

physiologisch aktivieren, bleiben langfristig besser haften als neutrale Stimuli.

Der Unterschied ist sowohl bei der Clusterliste signifikant ($U = 70$, $U_{\text{krit.}} = 87$ $p = .01$, einseitige Fragestellung) als auch bei der Willkürliste ($U = 65$, $U_{\text{krit.}} = 82$ $p = .01$). Die Nullhypothese kann zurückgewiesen werden.

Hypothese 7 (Vergleich Inc.K.Willkür - unabs.Willkür)

Die beiden Bedingungen des unabsichtlichen Lernens der Willkürliste (mit oder ohne Störaufgabe) unterscheiden sich hoch signifikant in ihrem Anteil an 'e erinnert'. Bei der Bedingung, bei der der Positionseffekt vermieden wurde, lag dieser Anteil wesentlich höher ($z = 2.5$ $p = .0062$, einseitige Fragestellung). Hypothese 7 trifft für das emotionale Erregungspotential, das aus den physiologischen Daten gewonnen wird, zu, was bei den Ergebnissen aufgrund des Gruppenurteils nicht der Fall war.

Hypothese 8 (Vergleich Frauen - Männer)

Die Werte der Frauen wurden mit denen der Männer über alle Bedingungen und innerhalb der verschiedenen Bedingungen verglichen. Die Frauen unterscheiden sich nicht von den Männern in ihrem Anteil an 'e erinnert'. Das Geschlecht spielt also für den Einfluß des emotionalen Erregungspotentials keine Rolle.

Vergleich der verschiedenen Intensitätsgrade

Von den 33 auswertbaren Wörtern fielen 11 in die n-Kategorie und 22 in die e-Kategorie. Von diesen 22 Wörtern waren durchschnittlich 2.4 Wörter stark erregend, 8 mittel erregend und 11.7 schwach erregend. Die mittlere Anzahl der Wörter in den einzelnen Intensitätsgraden und ihre absolute und relative Reproduktionshäufigkeit (Mittelwert und Prozent) unter den einzelnen Versuchsbedingungen sind in Tabelle 36 zusammengefaßt.

	stark erregend			mittel erregend			schwach erregend		
	>75		erinnert	75-25		erinnert	>25*		erinnert
	$\bar{X}$	$\bar{X}$	%	$\bar{X}$	$\bar{X}$	%	$\bar{X}$	$\bar{X}$	%
abs.Cl.	2.11	1.39	65.79%	8.50	4.56	53.59%	10.76	5.95	55.15%
Frauen:	2.00	1.56	77.78%	7.11	3.89	54.68%	11.11	6.89	62.00%
Männer:	2.22	1.22	55.00%	9.89	5.22	52.81%	10.44	5.00	47.87%
unabs.Cl.	3.00	2.05	68.33%	8.70	3.65	41.95%	10.95	4.20	38.36%
Frauen:	2.60	1.80	69.23%	6.70	2.90	43.28%	11.80	5.20	44.07%
Männer:	3.40	2.30	67.65%	10.70	4.40	41.12%	10.10	3.20	31.68%
abs.Wi.	2.35	1.18	50.00%	8.47	3.00	35.42%	11.24	4.41	39.27%
Frauen:	1.56	1.11	71.43%	8.22	3.00	36.49%	10.89	4.78	43.88%
Männer:	3.25	1.25	38.46%	8.75	3.00	34.29%	11.63	4.00	34.41%
unabs.Wi.	2.11	1.16	55.00%	7.26	2.05	28.26%	12.32	2.68	21.79%
Frauen:	2.00	1.00	50.00%	5.56	1.56	28.00%	11.22	2.89	25.74%
Männer:	2.20	1.30	59.09%	8.80	2.50	28.41%	13.30	2.50	18.80%
Inc.K.W.	2.35	1.04	44.41%	7.74	2.48	32.02%	13.87	4.43	31.97%
Frauen:	2.33	1.04	51.85%	8.25	2.27	32.32%	13.42	4.33	32.30%
Männer:	2.36	.91	38.46%	7.18	2.27	31.65%	14.36	4.55	31.65%
LT Cl.	2.56	.88	34.15%	8.00	2.25	28.13%	11.50	2.69	23.37%
Frauen:	2.38	.50	21.05%	6.00	1.63	27.08%	10.13	3.25	32.10%
Männer:	2.75	1.25	45.45%	10.00	2.88	28.75%	12.88	2.13	16.50%
LT Wil.	2.06	.88	42.42%	7.00	1.06	15.18%	10.50	1.56	14.88%
Frauen:	2.00	1.14	57.14%	6.43	1.29	20.00%	11.43	1.86	16.25%
Männer:	2.11	.67	31.58%	7.44	.89	11.94%	9.78	1.33	13.64%

Tabelle 36: Mittlere Anzahl der Wörter in den einzelnen Intensitätsgraden (PHY) und ihre absolute und relative Reproduktionshäufigkeit.

Die Friedman-Rangvarianzanalyse erbrachte für die Gesamtgruppen ($\chi^2_r = 19.29$; FG = 3; $p < .001$) und für die Gruppen getrennt nach Geschlecht ($\chi^2_r = 32.72$; FG = 3; $P < .001$) einen signifikanten Unterschied zwischen den vier Intensitätsstufen (incl. neutrale Wörter).

In der Gruppe mit den höchsten Ausschlägen wurden prozentual am häufigsten Wörter erinnert (einzige Ausnahme: Frauen LT Clusterliste). Von den neutralen Wörtern (kein Ausschlag) wurden prozentual am wenigsten Wörter erinnert (einzige Aus-

* ohne die neutralen Wörter, deren Werte in Tabelle 33 enthalten sind.

nahme: Männer unabs. Willkürliste; wegen des Positionseffekts). Das Bild bzgl. der mittel und schwach erregenden Wörter ist nicht einheitlich, teilweise ist die Behaltensleistung bei den mittel erregenden relativ höher als bei den schwach erregenden, teilweise ist es umgekehrt. Die Differenz zwischen den stark erregenden Wörtern und den mittlererregenden Wörtern erweist sich als signifikant, ebenso die Differenz zwischen den schwach erregenden und den neutralen Wörtern (Vorzeichen-test beide Male $p = .008$ für die Gesamtgruppen; $p = .001$ für die Einzelgruppen). Zwischen den mittel und schwach erregenden Wörtern besteht kein signifikanter Unterschied. Die vorliegenden Daten lassen eine positiv monotone Beziehung zwischen Erregungsintensität und Behaltensleistung vermuten. Eine Generalisierung ist bei diesen Daten allerdings nicht möglich, weil die hervorgerufene Erregung im Experiment eine bestimmte Höhe nicht überschritt, weshalb keine Aussagen über die Verhältnisse bei extrem starken Erregungen gemacht werden können.

6.2.3. Emotionales Erregungspotential aufgrund der subjektiven Einschätzung

Subjektive Einschätzungen lagen von 71 Vpn (36 Männern, 35 Frauen) vor. Einige Vpn konnten aus Zeitgründen daran nicht teilnehmen.

Tabelle 37 gibt die Daten der subjektiven Einschätzung (SE) an. Enthalten ist die Anzahl der erregend und der neutral beurteilten Wörter sowie ihre absolute und relative Reproduktionshäufigkeit.

Wie schon bei den physiologischen Daten wurde auch hier die Signifikanz der Differenzen der Prozentwerte zwischen E_{SE} erinnert und N_{SE} erinnert mit Hilfe des kritischen Bruchs und mit dem Wilcoxontest berechnet. Die Werte des kritischen Bruchs können Tabelle 38 entnommen werden.

	Vpn	E _{SE}	erinnert		N _{SE}	erinnert	
		$\bar{X}$	$\bar{X}$	%	$\bar{X}$	$\bar{X}$	%
abs.Cluster:	3	16.67	7.33	44.00%	23.33	8.33	35.71%
Frauen:	1	9.00	6.00	66.67%	31.00	15.00	48.39%
Männer:	2	20.50	8.00	39.02%	19.50	5.00	25.64%
unabs.Cluster:	15	21.73	9.73	44.79%	18.07	3.80	21.03%
Frauen:	7	23.29	11.00	47.24%	16.43	3.14	19.13%
Männer:	8	20.38	8.63	42.33%	19.50	4.38	22.44%
abs.Willkür:	5	29.00	9.20	31.72%	11.00	2.80	25.45%
Frauen:	3	29.33	10.00	34.09%	10.67	3.33	31.25%
Männer:	2	28.5	8.00	28.07%	11.50	2.00	17.39%
unabs.Willkür:	19	30.74	8.11	26.37%	9.26	1.47	15.91%
Frauen:	9	27.56	8.11	29.44%	12.44	1.44	11.61%
Männer:	10	33.60	8.10	24.11%	6.40	1.50	23.44%
Inc.K.Willkür:	18	29.83	8.67	29.05%	10.17	1.72	16.94%
Frauen:	10	29.30	9.00	30.72%	10.70	1.90	17.76%
Männer:	8	30.50	8.25	27.05%	9.50	1.50	15.79%
LT Cluster:	7	26.43	5.86	22.16%	13.57	1.43	10.53%
Frauen:	2	19.50	6.50	33.33%	20.50	2.00	9.76%
Männer:	5	29.20	5.60	19.18%	10.80	1.20	11.11%
LT Willkür:	4	31.00	5.00	16.13%	9.00	.75	8.33%
Frauen:	3	30.00	5.33	17.78%	10.00	1.00	10.00%
Männer:	1	34.00	4.00	11.76%	6.00	-	-

Tabelle 37: Anzahl erregend und neutral beurteilter Wörter mit den zugehörigen absoluten und relativen Reproduktionshäufigkeiten.

	CR	df	p
abs.Cluster:	.915	118	n.s.
Frauen:	1.010	38	n.s.
Männer:	1.294	78	n.s.
unabs.Cluster:	6.416	595	.0005
Frauen:	4.726	238	.0005
Männer:	3.625	276	.0005
abs.Willkür:	.891	198	n.s.
Frauen:	.295	118	n.s.
Männer:	1.794	78	.05
unabs.Willkür:	3.164	758	.005
Frauen:	4.257	358	.0005
Männer:	.116	398	n.s.

	CR	df	p
Inc.K.Willkür:	3.567	718	.0005
Frauen:	2.834	398	.005
Männer:	2.226	318	.025
LT Cluster:	2.651	278	.005
Frauen:	2.661	78	.005
Männer:	1.501	198	n.s.
LT Willkür:	1.376	158	n.s.
Frauen:	1.144	118	n.s.

Tabelle 38 (Forts.): CR-Werte der einzelnen Versuchsgruppen.

Nur vier der sieben Versuchsgruppen unterscheiden sich signifikant. Von ihren emotional eingeschätzten Wörtern reproduzieren diese Vpn prozentual mehr als von ihren neutral beurteilten Wörtern. Die drei Gruppen, bei denen sich kein Unterschied zeigt (abs.Cluster, abs.Willkür*, LT Willkür), sind hier die kleinsten Stichproben (nur 3 - 5 Vpn), deshalb sollte man mit einer Interpretation vorsichtig sein. Bei der nach Geschlecht getrennten Analyse ergibt sich zusätzlich für die männlichen Vpn bei der Versuchsbedingung 'unabs.Willkür, LT Cluster' kein Unterschied mehr.

Der Wilcoxontest konnte bei drei Versuchsgruppen (abs.Cluster, abs.Willkür, LT Willkür) nicht angewendet werden, da die Stichproben zu wenige Vpn umfaßten. Der Wilcoxontest bestätigt für die vier auswertbaren Gruppen die Werte des kritischen Bruchs. Der Unterschied zwischen E_{SE} -erinnert und N_{SE} -erinnert ist bei den Gruppen 'Inc.K.Willkür, unabs.Willkür, unabs.Cluster' (jeweils $p < .005$) und 'LT Cluster' ($p = .01$) signifikant.

Bei der Bedingung 'abs.Willkür' konnte der Vorzeichentest durchgeführt werden, der die Differenz als nicht signifikant auswies.

Bezogen auf Hypothese 1 läßt sich feststellen, daß die Ergebnisse aufgrund der subjektiven Einschätzung nicht mehr so

* Dieses Ergebnis wird nur durch die Daten der weiblichen Vpn verursacht; die männlichen Vpn weisen einen signifikanten Unterschied auf.

deutlich ausfallen, wie aufgrund des Gruppenurteils oder der physiologischen Daten.

Nun sollen die verschiedenen Versuchsbedingungen unter dem Aspekt des Anteils der emotional eingeschätzten Wörter an der Gedächtnisleistung (ohne Übungswörter) betrachtet werden. In Tabelle 39 stehen die prozentualen Werte für den Anteil E_{SE} -erinnert an der Gedächtnisleistung.

	Vpn	Anteil E_{SE}		Vpn	Anteil E_{SE}
abs.Cluster:	3	55.20%	Frauen:	1	57.94%
			Männer:	2	60.35%
unabs.Cluster:	15	68.05%	Frauen:	7	71.18%
			Männer:	8	65.35%
abs.Willkür:	5	55.48%	Frauen:	3	52.17%
			Männer:	2	61.75%
unabs.Willkür:	19	62.37	Frauen:	9	71.72%
			Männer:	10	50.70%
Inc.K.Willkür:	18	63.17%	Frauen:	10	63.37%
			Männer:	8	63.14%
LT Cluster:	7	67.79%	Frauen:	2	77.35%
			Männer:	5	63.32%
LT Willkür:	4	65.94%	Frauen:	3	64.00%
			Männer:	1	(100.00%)

Tabelle 39: Prozentualer Anteil E_{SE} -erinnertes Wörter an der Gedächtnisleistung.

Als Verfahren der Signifikanzprüfung wurde wieder der U-Test von Mann-Whitney herangezogen.

Hypothese 4 (Vergleich abs. - unabs.)

Absichtliches und unabsichtliches Lernen unterscheiden sich signifikant, ohne Kontrolle des Recency-Effekts ergibt sich der Wert $z = 1.876$ ($p = .03$); mit Kontrolle: $z = 2.208$ ($p = .013$). Beim absichtlichen Lernen ist der Anteil der emotional eingeschätzten Wörter an der Reproduktion geringer als beim unabsichtlichen Lernen. Eine mögliche Erklärung wäre, daß die Vpn beim absichtlichen Lernen weniger auf die Emotionalität der Wörter achten, weil sie ihre Konzentration auf die Lernaufgabe richten.

Für den Fall der subjektiven Einschätzung trifft Hypothese 4 zu, während sie bei den auf dem Gruppenurteil basierenden Daten und den physiologischen Daten nicht bestätigt werden konnte.

Bei allen anderen Variablen (Art der Liste, Länge des Retentionsintervalls, Störaufgabe, Geschlecht) zeigten sich keine signifikanten Unterschiede (über alle Bedingungen und innerhalb der einzelnen Bedingungen). Das bedeutet, daß die Nullhypothese von Hypothese 5, 6, 7, 8 beibehalten werden muß.

Unter den Ergebnissen dieses Abschnitts fällt vor allem auf, daß im Gegensatz zur Gruppenurteils-Analyse und den physiologischen Daten keine signifikante Differenz im Anteil emotional erregender Wörter zwischen Langzeitgedächtnis und Kurzzeitgedächtnis besteht. Bei der Langzeitgedächtnisprüfung werden prozentual nicht mehr emotionale Wörter erinnert als bei der Kurzzeitgedächtnisprüfung.

Vergleich der verschiedenen Intensitätsgrade

Tabelle 40 enthält die mittlere Anzahl der Wörter in den einzelnen Intensitätsgraden und ihre absolute und relative Reproduktionshäufigkeit (Mittelwert und Prozent).

Die Friedman Rangvarianzanalyse ergibt einen signifikanten Unterschied in der Reproduktionswahrscheinlichkeit der einzelnen Intensitätsgrade. Für die Daten der sieben Gesamtgruppen errechnet sich $\chi^2_r = 13.286$ (FG = 3; $p < .01$) und für die Daten der 14 Einzelgruppen $\chi^2_r = 25.76$ (FG = 3; $p < .001$). Die am stärksten erregend beurteilten Wörter weisen die höchste relative Reproduktionshäufigkeit auf.* Zwischen neutralen und schwach erregenden Wörtern gibt es keine Differenz (mit dem Vorzeichentest geprüft). Die mittel beurteilten Daten (der Einzelgruppen) unterscheiden sich diesmal signifikant von den schwach eingeschätzten Werten (Vorzeichentest: $p = .001$).

* Der Unterschied zu den anderen Intensitätsgraden ist auf dem 1%-Niveau signifikant (Vorzeichentest).

	stark erregend			mittel erregend			schwach erregend		
	>75 X	erinnert X	%	75-25 X	erinnert X	%	< 25* X	erinnert X	%
abs.Cl.	3.33	2.00	60.00%	8.67	4.33	50.00%	4.67	1.00	21.43%
Frauen:	3.00	2.00	66.67%	6.00	4.00	66.67%	-	-	-
Männer:	3.50	2.00	57.14%	10.00	4.50	45.00%	7.00	1.50	21.43%
unabs.Cl.	6.44	4.11	63.79%	13.11	5.56	42.37%	6.56	1.33	20.34%
Frauen:	8.50	5.50	64.71%	14.50	7.00	48.28%	5.25	1.25	23.81%
Männer:	4.80	3.00	62.50%	12.00	4.40	36.37%	7.60	1.40	18.42%
abs.Wi.	11.00	4.20	38.18%	13.40	3.20	23.88%	4.60	1.80	39.13%
Frauen:	11.33	5.00	44.12%	15.33	4.00	26.09%	2.67	1.00	37.50%
Männer:	10.50	3.00	28.57%	10.50	2.00	19.05%	7.50	3.00	40.00%
unabs.Wi.	6.95	2.79	40.15%	17.00	4.21	24.77%	6.79	1.11	16.28%
Frauen:	5.78	2.56	44.23%	14.56	4.33	29.77%	7.22	1.22	16.92%
Männer:	8.00	3.00	37.50%	19.20	4.10	21.35%	6.40	1.00	15.63%
Inc.K.W.	11.06	3.90	34.17%	14.61	4.00	27.38%	4.17	.83	20.00%
Frauen:	11.50	3.90	33.91%	14.20	4.30	30.28%	3.60	.60	16.67%
Männer:	10.50	3.63	34.52%	15.13	3.63	23.97%	4.88	1.13	23.08%
LT Cl.	5.43	2.57	47.37%	10.86	2.43	22.37%	10.14	.86	8.48%
Frauen:	6.00	3.00	50.00%	8.00	3.00	37.50%	5.50	.50	9.09%
Männer:	5.20	2.40	46.15%	12.00	2.20	18.33%	12.00	1.00	8.33%
LT Wil.	6.50	2.00	30.77%	13.75	2.50	18.18%	10.75	.25	2.33%
Frauen:	4.00	1.67	55.56%	13.33	3.00	22.50%	12.67	.33	2.63%
Männer:	14.00	3.00	21.43%	15.00	1.00	6.67%	5.00	-	-

Tabelle 40: Mittlere Anzahl der Wörter in den einzelnen Intensitätsgraden (SE) und ihre absolute und relative Reproduktionshäufigkeit.

6.2.4. Vergleich der drei Arten des emotionalen Erregungspotentials hinsichtlich der Behaltensleistung

In der folgenden Tabelle 41 sind die Werte der drei Maße (prozentualer Anteil emotional erregender Wörter an der Gedächtnisleistung) zusammengestellt. Die Daten sind in etwa geordnet nach der Größe des prozentualen Anteils.

* ohne die neutralen Wörter, deren Werte in Tabelle 37 aufgeführt sind.

	PHY	G*	SE
LT Cluster:	84.86%	84.31%	67.79%
LT Willkür:	82.59%	75.38%	65.94%
Inc.K.Willkür:	75.02%	62.62%	63.17%
abs.Willkür:	74.95%	61.82%	55.48%
unabs.Cluster:	68.32%	67.61%	68.05%
abs.Cluster:	63.53%	62.54%	55.20%
unabs.Willkür:	60.35%	65.19%	62.37%

	PHY	G	SE
LT Cluster Männer:	90.91%	80.77%	63.32%
LT Willkür Männer:	86.96%	85.19%	(100.00%) **
LT Cluster Frauen:	83.94%	88.00%	77.35%
abs.Willkür Frauen:	80.33%	61.82%	52.17%
LT Willkür Frauen:	79.08%	68.42%	64.00%
Inc.K.Willkür Frauen:	77.77%	58.93%	63.17%
Inc.K.Willkür Männer:	72.44%	66.67%	63.14%
unabs.Cluster Frauen:	71.39%	66.41%	71.18%
abs.Cluster Männer:	70.68%	64.29%	60.35%
abs.Willkür Männer:	66.43%	65.52%	61.75%
unabs.Cluster Männer:	65.13%	68.91%	65.35%
unabs.Willkür Frauen:	62.26%	71.25%	71.72%
abs.Cluster Frauen:	61.77%	61.14%	(57.72%) **
unabs.Willkür Männer:	56.78%	61.00%	50.70%

Tabelle 41: Prozentualer Anteil emotional erregender Wörter an der Gedächtnisleistung nach Gruppenurteil (G), physiologischen Daten (PHY) und subjektiver Einschätzung (SE).

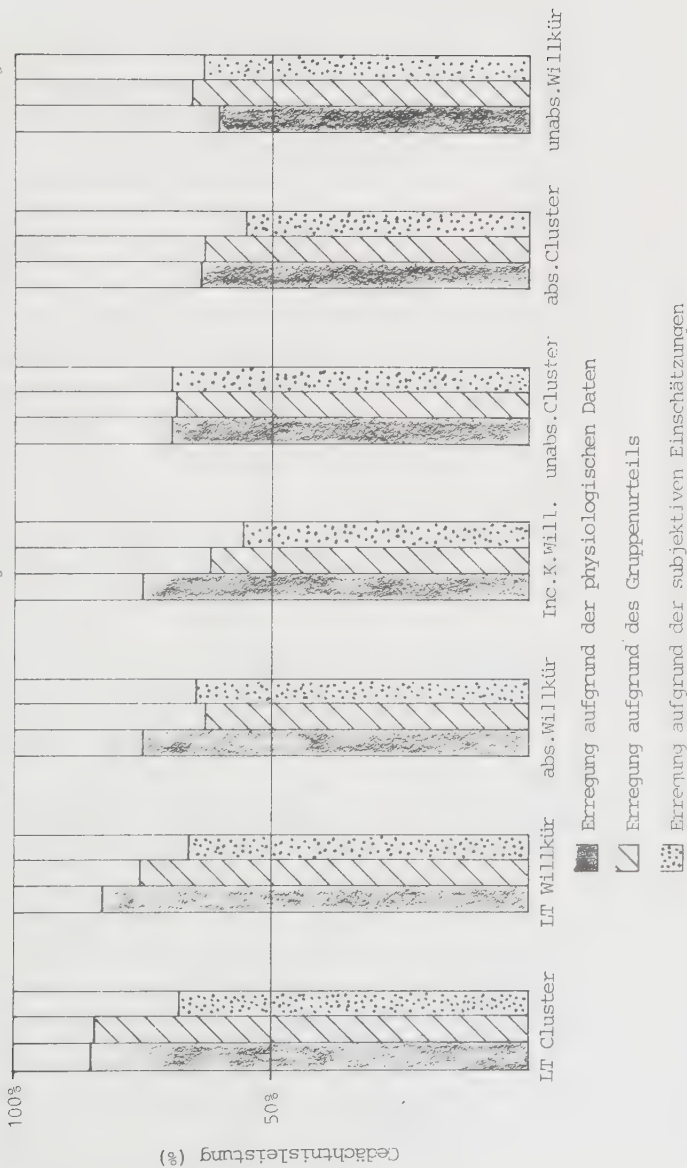
Veranschaulicht werden die Werte der Gesamtgruppen in Abb. 26.

Zur Prüfung von Hypothese 2 und 3 wurde der Vorzeichentest (einseitige Fragestellung) herangezogen. (abhängige Stichproben).

* Hier handelt es sich nicht um die Mittelwerte, wie auf S. 201 und 203 dieser Arbeit, sondern um Prozentwerte der Gruppe. Es wurden nur die Vpn gewertet, von denen physiologische Daten vorlagen.

** Die Daten stammen nur von einer Vp.

Abb. 26: Anteil emotional erregender Wörter an der Gedächtnisleistung.



Rest der jeweiligen Säule enthält die neutralen Wörter

Hypothese 2 (Vergleich normativ - ipsativ)

Zuerst erfolgt der Vergleich des Gruppenurteils mit den physiologischen Daten. Daran schließt sich der Vergleich des Gruppenurteils mit den subjektiven Einschätzungen an.

a) Vergleich: Gruppenurteil- physiologische Daten

Die Hypothese besagt, daß der Anteil emotional erregender Wörter an der Gedächtnisleistung bei der Zugrundelegung der physiologischen Aktivierung höher ausfällt als aufgrund des a priori Gruppenurteils. Es wird angenommen, daß die physiologische Erregung das 'wahre' Erregungspotential von Wörtern zuverlässiger mißt. Im Gegensatz zu den physiologischen Werten (ipsative Auswertung) werden beim Gruppenurteil spezifische Unterschiede zwischen den Vpn nicht berücksichtigt.

Die Analyse der Daten führt zu einer Bestätigung der Hypothese 2. Der Vorzeichentest erbrachte über alle Vpn den Wert $z = 2.69$. Die Wahrscheinlichkeit, daß dieses Ergebnis zufällig ist, beträgt $p = .0036$.

Bei der getrennten Betrachtung der einzelnen Gruppen ergibt sich folgendes Bild. Bei der Willkürliste ($p = .0436$) und bei der Clusterliste ($p = .0228$) ist der Anteil emotionaler Wörter an der Reproduktion auf der Grundlage der physiologischen Werte signifikant ($p < .05$) höher als der Anteil auf der Grundlage des Gruppenurteils.

Die getrennte Analyse der Daten der weiblichen und der männlichen Vpn weist nur für die Frauen ($p = .0096$) einen signifikanten Unterschied auf. Bei den Männern wird der Unterschied bis auf die Bedingung 'LT Cluster' ($p = .016$) nicht signifikant ($p = .0934$).

Die Unterschiede sind sowohl hinsichtlich des Kurzzeitgedächtnisses ($p = .0409$) als auch für das Langzeitgedächtnis ($p = .008$) signifikant. Das gleiche gilt für die absichtliche und unabsichtliche (incl. LT) Lernbedingung ($p = .0301$ bzw. $.0307$). Nur für die unabsichtliche Lernbedingung ST ergibt sich kein Unterschied, was auf die Kontamination der Daten durch den Positionseffekt zurückzuführen ist.

b) Vergleich: Gruppenurteil - subjektive Einschätzungen
Über alle Vpn besteht kein Unterschied. Der Anteil emotional erregender Wörter an der Gedächtnisleistung ist ähnlich hoch, gleichgültig ob die Emotionalität der Wörter durch das Gruppenurteil oder die subjektive Einschätzung bestimmt wird. Eine leichte Tendenz ist erkennbar, wonach bei den Bedingungen 'längerfristiges Behalten' und 'absichtliches Lernen' entgegen der Hypothese die Gruppenwerte besser ausfallen. Die Nullhypothese kann hier nicht zurückgewiesen werden.

Hypothese 2 trifft nur für den Fall zu, daß die ipsativen Werte auf physiologischen Daten basieren.

Hypothese 3 postuliert, daß die physiologische Reaktion das Erregungspotential besser erfaßt als die Beurteilung der Vp dies vermag. Die Erregungen können von den Vpn nicht differenziert genug wahrgenommen werden, da sie sich teilweise im unmerklichen Bereich abspielen.

Über alle Vpn bestätigt der Vorzeichentest unsere Hypothese, nach der die Werte auf der Basis der physiologischen Daten höher ausfallen als auf der Grundlage der subjektiven Beurteilungen ($z = 2.0158$ $p = .0202$).

Bei der Betrachtung der Einzelgruppen erweist sich der Unterschied nur bei der Clusterliste als signifikant ($p = .047$), ebenso beim absichtlichen Lernen ($p = .035$) und beim Langzeitgedächtnis ($p = .008$).

Beim unabsichtlichen Lernen der Willkürliste verzerrt der Recency-Effekt die Daten, so daß kein signifikanter Unterschied entsteht.

6.3. Emotionales Erregungspotential und Rekognition

Im folgenden werden die Ergebnisse der Rekognition wiedergegeben. Auf der linken Seite stehen in Großbuchstaben die Übungs- und Testwörter, daneben die zu diesen Wörtern gehörigen Distraktoren. Die Zahl hinter dem Wort gibt an, wie oft dieses Wort angestrichen wurde. Bei beiden Versuchsgruppen war die Anzahl der Vpn 16.

Den Vpn war freigestellt, wieviele Wörter sie ankreuzen wollten. Fast alle Vpn markierten weniger Wörter als sie bei der Wortdarbietung gesehen hatten.

Bei der Willkürliste wurden von jeder Vp durchschnittlich 27 Items gekennzeichnet, davon entfielen 16 auf die Wörter der Lernliste (14 Testwörter, 2 Übungswörter) und 11 auf die Distraktoren.

Bei der Clusterliste kreuzte jede Vp durchschnittlich 40 Wörter an. 21 stammten aus der Wortdarbietung (19 Testwörter, 2 Übungswörter). 19 Wörter wurden fälschlicherweise angestrichen.

Von den Störitems wurden hauptsächlich die den Testwörtern bedeutungsähnlichen Wörter gewählt, seltener die phonetisch ähnlichen. Leider ließ sich bei einigen phonetischen Distraktoren eine inhaltliche Beziehung zu anderen Test- oder Übungswörtern nicht vermeiden. Z.B. weist das Störwort 'Kuh' (Testwort 'Kuß') einen bedeutungsmäßigen Zusammenhang mit dem Übungswort 'Tier' auf.

Es fiel auf, daß bei der Langzeitgedächtnisprüfung eher konkrete Items wiedergegeben wurden als abstrakte. Anstelle des Wortes 'Farbe' wurde von einigen Vpn das Wort 'rot' angestrichen. Veränderungen vom Substantiv zum Verb traten bei den Wörtern Luft (atmen), Sprache (sprechen) und Tränen (weinen) auf.

Rekognitionsliste (Willkürliste)

Übungswörter

FARBE 5,	Farce -, färben -, rot 3
NEUN 2,	nein -, Zahl -, zehn, rechnen 1
HAND 3,	Wand -, Arm -, Fuß -, Finger -
SCHREIBEN 7,	schreien 3, Schrift 1, lesen 4
BEGRIFF 7,	Beginn 1, Ausdruck 3, begreifen 2
VERKEHR 10,	verkehrt -, Umgang -, Koitus 2, Straße 1

Testwörter

ABBILDUNG 3,	Abbitte -, abbinden-, Bild 1, Foto -
ABGRUND 7,	Abguß -, Grund -, Tiefe 1, fallen -
AMPUTATION 8,	Amplitude 1, Ampulle -, Abtrennung 1, Kastration 3
AUSGEGLICHEN 3,	ausgelassen -, ruhig 1, nervös 2, heiter 1
BORDELL 10,	Boden -, bohren 1, Puff 1, Dime 3
BUCHSTABE 2,	Buch 1, Silbe -, Wort 1, Paragraph -
FERIEN 7,	Ferne -, Urlaub 1, Arbeit 1, frei 1
FOLTER 8,	Folge 1, folgern -, Tortur -, quälen 1
GEBORGENHEIT 6,	Gebrechlichkeit -, geborgen 1, Wärme -, Zärtlichkeit 1
GEGEND 2,	gegen -, Gebiet -, Umgebung -, Ort -
GEISTESKRANK 3,	Geistlicher -, nervenkrank 1, schizophren 3, Idiot 4
GELD 5,	gelb 1, Vermögen 1, Mark -, Kapital 1
HEIZUNG 4,	Zeitung -, heimlich -, Ofen 3, heiß -
KANTE 6,	Tante -, Karte -, Rand 1, Ecke -
KIRCHE 3,	Kirsche -, Religion 1, Glaube -, Gott 1
KOTZEN 8,	Kot 3, köpfen -, rotzen 1, erbrechen 3
KRIEG 8,	kriechen 1, Kampf 1, Militär -, Politik 1
LAMPE 3,	lange -, Lamm -, Leuchter -, Licht 2
LEICHE 10,	leicht -, Toter 1, Sarg 3, sterben 2
LUFT 6,	Lust 1, Lift 1, Lunge -, atmen 4
MATERIAL 2,	Materie 2, Stoff 1, Holz 2, Eisen -
MUSIK 6,	Muster -, Lied -, Konzert -, singen -
ORGASMUS 12,	Organismus 3, Organ 1, Höhepunkt 1, Befriedigung 2
PRÜFUNG 7,	Prüderie -, prude 1, Prügel 1, Examen 2
PSYCHOLOGIE 12,	Physiologie 1, Psychiatrie 2, Psychologe 1, Seele 2

RAD 5,	Rat 1, Fahrrad 2, Reifen -, Auto 2
RECHTECKIG 5,	rechtzeitig -, Recht 1, viereckig 2, rund 1
REGAL 5,	Regel 2, Wandgestell -, Bord -, Schrank 1
REGENSCHIRM 3,	Regentschaft -, Schirm 2, Regen 4, Wetter -
SAU 9,	Bau -, klauen 1, Schwein 3, Kuh 1
SCHWANGER 13,	schwanken -, schwul 5, trächtig -, Kind 2
SEHEN 2,	gehen -, schauen 1, Auge 1, Wahrnehmung 4
SPRACHE 2,	Spucke 1, spröde 1, sprechen 4, reden -
STIEFEL 2,	Spiegel -, Stiel -, Schuhe 1, Winter -
STREIT 5,	Streik -, Streich -, Zwist -, Streß 7
TRÄNEN 6,	Trainer -, träge -, weinen 4, Blut 2
TREPPE 3,	Trippler -, Treue 3, Stufe -, Stiege -
UNTERSCHIED 1,	Unterschicht -, Differenz -, Ungleichheit 1, unterscheiden -
VERBRECHEN 8,	verbrennen 1, Gebrechen 1, Kriminalität 5, Verbrecher 5
WOHNUNG 2,	wohlgemut -, Wohlstand 1, Heim -, wohnen -

Rekognitionsliste (Clusterliste)

Übungswörter

GEGENSTAND 9,	Gegenwart 2, Ding 6, Objekt 6
TIER 3,	Tisch 5, Katze 2, Hund 2
DURCHMESSER 7,	Messer 3, Gerade 1, Dicke -
LUFT 8,	Lust 5, Atem -, Lunge -
BLEISTIFT 8,	Blei 1, Farbstift 1, Radiergummi 2
STRASSE 4,	Strafe 2, Weg -, Verkehr 4, Auto 1

Testwörter

BEERDIGUNG 14,	Beendigung -, Beleidigung -, Erde 2, Friedhof 8
BEGRIFF 6,	Beginn 2, Griff -, Ausdruck -, Bedeutung 3
BUCHSTABE 11,	Buch 1, Silbe 2, Vokal 2, Paragraph -
BUSEN 14,	Buße 3, Schmusen 5, Brust 3, Rücken -
DACHSTUHL 5,	Dachs -, Stuhl 5, Dach 2, Speicher 2
DURCHFALL 10,	Durchfahrt -, Stuhlgang 3, Verstopfung 2, Mißerfolg 1
EISEN 4,	Eis -, Metall 4, Stahl 3, Beton 1
FUSS 1,	Faß -, Bein -, Zehe -, Sohle -
GABEL 6,	Galgen -, Galle -, Löffel -, Schere -
GESCHIRR 5,	Geschick 1, Geschlecht 3, Hausrat -, Werkzeug 4
GESTANK 9,	Gestalt 1, Mief 1, Geruch 2, Duft 2
GOTT 9,	Gold 1, Christus 1, Glaube 3, Religion 6
HAND 5,	Wand 3, Arm -, Finger 1, Körper 1
HAUS 5,	Maus -, Gebäude 1, Heim 1, Garten 2
HÖLLE 5,	Hülle 1, Satan 1, Teufel 1, Himmel 4
HOLZ 8,	Hof 1, Klotz 1, Baum 1, Späne -
KANNE 3,	Kante 1, Tanne -, Krug 2, Eimer 3
KIRCHE 2,	Kirsche 1, Messe -, Pfarrer 2, Altar 2
KLOPAPIER 7,	Papier 1, Kloacke -, Klosett 3, Toilette 3
KOT 9	Kost -, Dreck -, Schmutz 1, Darm 1
KUSS 12	Kuh 1, Kurs -, Mund 3, Lippen 4
LIEBE 8,	Lied -, Libido -, Gefühl 3, Freundschaft 3
MATERIAL 7	Materie 5, Matrikel 1, Unterlagen -, Belege -
ORGASMUS 16,	Organismus 1, Organ 1, Höhepunkt 4, Befriedigung 2
PENIS 14,	Pendel -, Pension -, Pensum -, Glied 3
PRÜFUNG 11,	Prüderie 2, Prügel -, Examen 5, Test 2

PSYCHOLOGIE 4,	Physiologie 1, Psychiater -, Psychologe 1, Seele 1
SARG 7,	Sage -, Toter 4, Grab 5, Leiche 4
SATZ 5,	Salz -, Syntax 1, Sprache 4, Grammatik 2
SCHORNSTEIN 5	Stein 2, Kamin 2, Schornsteinfeger -, Ruß -
STERBEN 12,	Stelle -, Trauer 6, Leiden 1, Siechtum -
STOFF -,	Stock -, Tuch -, Gewebe -, Rauschgift 1
STUDIUM 9,	Stunde -, Ausbildung 2, Student 6, Beruf 1
SÜNDE 8,	Sühne 1, Vergehen -, Beichte 1, Gebot 3
TELLER 3,	Tellur 1, Keller -, Tasse 1, Topf 1
TOD 10,	Ton 3, Ende 3, Verwesung 5, Leben 4
UNIVERSITÄT 9,	Universum 1, Hochschule 3, Wissenschaft 6, Bildung 3
WORT 8,	Wert 3, Hort -, Rede 1, Versprechen 2
ZÄRTLICHKEIT 13,	Zänkerei -, Liebkosung 3, Sexualität 7, Geborgenheit 3
ZIMMER 4,	Ziffer 1, Küche -, Raum 3, Wohnung 2

Bei einem Teil der Wörter zeigte sich eine große Diskrepanz zwischen dem Wiedererkennen und der freien Reproduktion. Diese Wörter wurden kaum reproduziert, aber von vielen Vpn auf der Rekognitionsliste wiedererkannt.

Bei der Clusterliste verzeichneten folgende Wörter eine starke Steigerung beim Wiedererkennen: die Wörter der Kategorie 'Toilette', außerdem die Wörter 'Buchstabe*', Wort, Material*, Holz, Sterben, Prüfung, Studium, Zärtlichkeit, Kuß, Gott und Sünde'. Die Wörter 'Material, Holz und Gott' wurde bei der freien Reproduktion nicht erinnert, aber dann in der Rekognitionsliste etwa von der Hälfte der Vpn wiedererkannt.

Bei der Willkürliste handelt es sich dabei um die Wörter: 'schwanger, Leiche, Sau, Folter, kotzen, Amputation, Verbrechen, Krieg, Ferien, Geld, Kante und Regal'.

Auch für das Wiedererkennen macht sich der Einfluß des emotionalen Erregungspotentials bemerkbar. Häufig wiedererkannt wurden vor allem die emotional erregenden Wörter.

Die Werte des Anteils emotional erregender Wörter an der Gedächtnisleistung (Wiedererkennen der Testwörter) lauten:

<u>nach Gruppenurteil:</u>	Clusterliste	66.79%
	Willkürliste	70.79%

nach physiologischen Daten:

	Clusterliste	60.25%
	Willkürliste	57.00%

* Diese Wörter wurden vermutlich aufgrund der Wirkung der Cluster besser wiedererkannt, da sie bei der Willkürliste nicht sehr viel besser erkannt (nur von 2 Vpn) als reproduziert wurden.

7. Zusammenfassung

In dieser Arbeit wurde der Versuch unternommen, den Einfluß des emotionalen Erregungspotentials von Wörtern auf das verbale Lernen zu ermitteln. Die bisherigen Untersuchungen in diesem Bereich erbrachten wechselnde Ergebnisse. Verursacht wurden diese vor allem durch eine Reihe von Mängeln und Problemen, die in der vorliegenden Untersuchung durch die Kontrolle von Störfaktoren und die Berücksichtigung mehrerer Operationalisierungen vermieden werden sollten. Das emotionale Erregungspotential wurde mit Hilfe des Gruppenurteils, der subjektiven Einschätzung und der physiologischen Messung (EDR) bestimmt.

Als wesentliches Ergebnis der Untersuchung ist festzuhalten, daß emotional erregende Wörter besser behalten werden als neutrale, unabhängig davon auf welche Art das emotionale Erregungspotential operationalisiert wird.

Dabei erwies sich die EDR als bester Indikator für die Behaltensleistung. Der Anteil der durch Gruppenurteil und subjektive Einschätzung als erregend beurteilten Wörter an der Gedächtnisleistung war geringer als aufgrund der physiologisch festgestellten Daten. Die bewußt wahrgenommene und beurteilte Erregung scheint also nicht so verhaltenswirksam zu sein wie die vegetative physiologische Aktivierung, die hier mit Hilfe der EDR gemessen wurde.

Die Lernbedingungen 'Art der Liste', 'Absichtlichkeit des Lernens' und das 'Geschlecht der Vp' haben vermutlich keinen Einfluß auf die Wirkung des emotionalen Erregungspotentials von Wörtern hinsichtlich der Gedächtnisleistung. Diese Bedingungen unterschieden sich nicht in der Höhe des Anteils emotional erregender Wörter an der Reproduktion.

Auch kurzfristig werden mehr emotional erregende Wörter reproduziert, wenn der Recency-Effekt ausgeschaltet ist. Die

Konsolidierungstheorie von WALKER & TARTE (1963), die eine Abrufhemmung stark erregender Wörter beim Kurzzeitgedächtnis postuliert, kann nicht bestätigt werden. Sie beruht vermutlich nur auf einem durch den Recency-Effekt hervorgerufenen Artefakt.

Langfristig bleiben im Gedächtnis vor allem die emotional erregenden Wörter (PHY, Gruppenurteil) haften. Der hohe Anteil erregender Wörter ist darauf zurückzuführen, daß bei der Langzeitgedächtnisprüfung kaum neutrale Wörter erinnert wurden.

Die Ergebnisse der Untersuchung zeigen, daß eine emotionale Erregung günstiger ist für das Einprägen von Stimuli.

Anhang

2	6,50000	1,00000	5,50000	7,40000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
3	15,00000	5,50000	7,40000	8,00000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
4	6,31200	1,00000	5,50000	7,40000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
5	5,50000	5,50000	7,40000	8,00000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
6	8,16700	4,50000	6,00000	7,40000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
7	12,00000	15,00000	7,40000	8,00000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
8	8,50000	5,33300	13,50000	7,40000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
9	6,14300	8,50000	7,00000	7,40000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
10	8,50000	4,00000	3,00000	8,00000	8,00000	12,20000	7,16700	6,66700	1,62500	4,00000	4,00000
11	7,00000	9,00000	7,66600	6,33300	5,50000	8,00000	11,50000	11,50000	11,50000	11,50000	11,50000
12	5,47000	5,50000	8,50000	7,43700	5,50000	8,00000	15,00000	15,00000	15,00000	15,00000	15,00000
13	8,00000	4,00000	5,50000	4,00000	5,33300	7,50000	10,00000	10,00000	10,00000	10,00000	10,00000
14	14,00000	10,00000	6,50000	8,53600	6,83000	10,71400	8,60000	8,60000	8,60000	8,60000	8,60000
15	11,50000	8,50000	7,30000	7,30000	6,50000	7,60000	12,00000	12,00000	12,00000	12,00000	12,00000
16	6,00000	3,71400	12,00000	5,00000	9,50000	14,00000	1,00000	1,00000	1,00000	1,00000	1,00000
17	13,00000	9,50000	16,00000	8,00000	11,50000	14,00000	1,00000	1,00000	1,00000	1,00000	1,00000
18	17,50000	22,00000	6,00000	10,00000	11,50000	14,00000	1,00000	1,00000	1,00000	1,00000	1,00000
19	9,16700	2,00000	6,00000	5,80000	6,66700	7,50000	8,33300	8,33300	8,33300	8,33300	8,33300
20	4,00000	3,33300	7,00000	9,57100	7,80000	14,20000	9,00000	9,00000	9,00000	9,00000	9,00000
21	1,00000	7,00000	7,22200	7,16700	7,80000	14,20000	9,00000	9,00000	9,00000	9,00000	9,00000
22	4,00000	7,00000	7,22200	7,16700	7,80000	14,20000	9,00000	9,00000	9,00000	9,00000	9,00000
23	6,66700	5,50000	2,00000	8,00000	8,33300	18,00000	10,33300	10,33300	10,33300	10,33300	10,33300
24	2,00000	5,50000	2,00000	8,00000	8,33300	18,00000	10,33300	10,33300	10,33300	10,33300	10,33300
25	5,27300	8,50000	8,33300	2,00000	7,50000	17,50000	7,00000	7,00000	7,00000	7,00000	7,00000
26	8,50000	8,50000	8,60000	8,27300	4,67000	17,50000	7,00000	7,00000	7,00000	7,00000	7,00000
27	6,50000	13,33300	13,30000	4,66700	11,00000	17,50000	7,00000	7,00000	7,00000	7,00000	7,00000
28	3,83300	8,00000	7,44700	7,25000	8,00000	16,66700	1,00000	1,00000	1,00000	1,00000	1,00000
29	17,00000	14,50000	4,00000	4,50000	15,50000	2,00000	8,00000	8,00000	8,00000	8,00000	8,00000
30	4,00000	15,31200	10,30000	2,70600	5,50000	7,83300	4,83300	4,83300	4,83300	4,83300	4,83300
31	6,13300	9,21400	6,85700	6,90000	8,71400	4,66700	8,00000	8,00000	8,00000	8,00000	8,00000
32	2,92300	5,67700	5,16700	6,90000	8,71400	4,66700	8,00000	8,00000	8,00000	8,00000	8,00000
33	5,85400	7,80000	3,00000	1,76200	6,50000	9,66700	5,30000	5,30000	5,30000	5,30000	5,30000
34	3,25000	7,25300	2,50000	2,50000	7,90000	7,66700	5,30000	5,30000	5,30000	5,30000	5,30000
35	10,12500	7,00000	6,50000	7,44400	6,33300	11,00000	4,75000	4,75000	4,75000	4,75000	4,75000
36	11,50000	8,50000	7,25000	9,00000	11,80000	8,00000	10,25000	10,25000	10,25000	10,25000	10,25000
37	3,00000	7,55000	5,50000	6,00000	6,00000	7,50000	8,83700	8,83700	8,83700	8,83700	8,83700
38	7,24400	4,00000	2,00000	6,80000	5,20000	7,50000	11,33300	11,33300	11,33300	11,33300	11,33300
39	11,25000	6,66700	9,50000	7,35700	8,50000	11,33300	11,33300	11,33300	11,33300	11,33300	11,33300
40	7,00100	8,14000	4,25000	6,82400	6,20000	4,00000	6,16700	6,16700	6,16700	6,16700	6,16700

Tabelle: Matrix der Distanz-Koeffizienten. Die 40 Wörter sind alphabetisch geordnet. (CLUSTERLISTE)

28	3.26000	8.50000	12.50000	6.81800	7.60000	12.66700	9.16700	9.75000	9.46700	9.56000
	9.33000	4.66700	12.75000	8.90000	4.71400	10.56000	8.50000	3.25000	3.80300	9.11100
29	6.80000	6.00000	9.80000	8.13000	6.14300	10.71400	8.20000	9.06000	17.50000	12.06000
	12.20000	12.50000	4.00000	1.40000	1.30000	13.00000	7.50000	15.0000	6.00000	8.00000
30	5.80000	11.3300	3.00000	8.16700	8.60000	3.00000	3.20000	7.50000	5.75000	2.66700
	8.50000	3.00000	3.00000	9.75000	3.66700	9.00000	3.0000	7.50000	5.00000	7.50000
	3.66700	8.25000	13.00000	8.86700	8.66700	6.00000	1.50000	6.0000	4.00000	7.50000
31	3.40000	7.66700	5.67000	3.00000	3.43300	6.00000	6.00000	5.33000	8.33300	7.50000
	6.00000	5.10000	10.20000	10.20000	2.50000	14.00000	11.75000	12.50000	4.75000	1.50000
	12.25000	10.60000	6.00000	9.20000	7.50000	9.50000	1.50000	10.0000	4.75000	5.50000
32	6.00000	7.87000	9.40000	8.76000	7.33300	3.50000	1.50000	10.33300	3.00000	7.50000
	9.00000	7.25000	11.00000	6.57100	10.0000	3.00000	1.00000	3.00000	3.00000	7.50000
	3.23300	4.00000	1.00000	6.57100	7.14300	3.00000	6.00000	6.33300	1.50000	2.66700
33	6.00000	2.50000	6.81200	4.60000	7.40000	9.71400	9.71400	6.00000	10.30000	10.90000
	9.33000	5.00000	9.55700	8.50000	7.45700	8.50000	10.83300	9.44400	7.78000	7.80000
	6.13400	8.25000	7.77800	5.68700	5.77800	3.50000	2.50000	7.27300	4.15700	9.00000
34	4.66700	7.00000	6.83300	6.66700	5.80000	17.00000	9.75000	9.75000	13.33300	13.00000
	5.50000	3.1400	12.33300	2.33300	23.00000	11.00000	11.00000	4.60000	7.00000	5.60000
	8.16700	8.42900	15.66700	10.12500	8.85700	9.50000	9.33300	5.33300	16.00000	5.00000
35	2.50000	7.30000	3.50000	7.90000	6.00000	14.66700	5.25000	5.25000	1.25000	1.75000
	6.13300	9.42900	5.00000	4.57100	7.40000	13.00000	1.57100	10.20000	4.00000	2.25000
	4.50000	8.25000	7.50000	7.50000	7.60000	6.40000	10.33300	10.0000	12.50000	3.00000
36	11.50000	7.00000	11.66700	11.50000	8.83300	7.42900	12.87500	8.12500	4.14300	7.83300
	3.74000	7.20000	5.00000	5.33300	5.37500	14.66700	4.60000	5.96900	9.00000	7.70000
	11.00000	5.07100	9.00000	7.05300	5.25000	7.66700	7.11100	2.91700	10.75000	7.00000
37	2.33300	7.00000	6.12500	4.22000	8.20000	8.33000	5.80000	4.63600	4.25000	8.42500
	8.07700	5.09000	7.30000	6.68400	4.50000	6.75000	5.00000	7.40000	5.28500	6.50000
	3.00000	5.57100	7.20000	6.68400	4.25000	6.75000	7.33300	7.75000	16.33300	11.90000
38	4.68600	1.56700	2.00000	6.66700	10.20000	12.50000	5.33300	13.33300	5.33300	5.60000
	7.33300	5.5000	12.00000	7.50000	10.20000	8.75000	4.88900	6.60000	1.60000	5.00000
39	3.00000	12.00000	11.44000	3.66700	11.33300	8.14400	4.66700	5.62500	9.60000	9.75000
	7.33000	1.250000	1.44000	3.66700	7.50000	9.00000	7.20000	7.10000	7.33000	8.75000
	9.00000	7.93300	5.20000	7.33300	6.00000	7.33300	7.00000	6.70000	6.40000	6.66700
	3.65200	1.93200	5.20000	4.29400	6.05300	6.82400	6.75000	7.63600	6.40000	6.66700
40	14.00000	2.00000	4.50000	9.50000	9.37500	8.27500	9.53300	7.63600	4.33300	4.00000
	1.00000	7.00000	4.50000	10.50000	4.60000	15.00000	9.00000	6.00000	12.50000	12.50000
	12.00000	6.60300	4.00000	7.00000	7.00000	8.00000	9.00000	9.66700	8.50000	2.50000
	5.00000	5.60300	7.00000	15.66700	7.42900	4.33300	6.85700	7.75000	10.33300	10.33300
	2.00000	8.20000	8.20000	3.75000	12.00000	12.00000	5.83300	7.75000	7.75000	7.75000

2	5,00000	6,00000	6,25000	3,75000	7,50000	11,67000	5,6000	9,00000	5,87000
3	6,67000	7,00000	4,40000	10,33000	5,50000	7,00000	4,50000	2,67000	
4	5,00000	3,60000	8,50000	3,00000	5,00000	6,00000	8,00000	5,00000	
5	5,17000	8,00000	5,50000	3,00000	5,00000	7,00000	4,50000		
6	4,00000	3,50000	2,33000	2,65000	3,00000	4,50000	6,00000		
7	7,00000	6,50000	2,00000	3,25000	4,17000	6,50000	8,00000		
8	5,00000	2,33000	3,00000	8,33000	3,50000	8,60000	3,25000		
9	8,00000	3,00000	2,50000	8,00000	6,75000	8,60000	4,00000		
10	7,25000	3,30000	3,00000	5,00000	7,00000	8,60000	4,00000		
11	14,50000	3,00000	8,00000	6,00000	5,75000	8,60000	4,00000		
12	3,00000	5,60000	5,00000	6,40000	7,00000	8,60000	4,00000		
13	3,50000	1,50000	5,00000	6,40000	5,75000	8,60000	4,00000		
14	3,80000	5,00000	7,00000	6,40000	5,75000	8,60000	4,00000		
15	6,00000	5,00000	10,00000	6,40000	5,75000	8,60000	4,00000		
16	6,27000	3,33000	6,67000	5,60000	7,00000	8,60000	4,00000		
17	10,25000	4,50000	3,00000	8,47000	5,50000	8,60000	4,00000		
18	8,75000	5,00000	4,60000	7,50000	7,00000	8,60000	4,00000		
19	5,33000	4,67000	6,50000	7,17000	7,00000	8,60000	4,00000		
20	2,57000	10,20000	2,00000	6,00000	5,67000	8,60000	4,00000		
21	8,50000	2,00000	7,33000	4,47000	4,00000	8,60000	4,00000		
22	6,00000	8,30000	5,98000	6,50000	8,00000	8,60000	4,00000		
23	6,12000	5,67000	3,80000	11,00000	5,33000	8,60000	4,00000		
24	6,50000	7,67000	5,80000	9,00000	6,00000	8,60000	4,00000		
25	5,17000	6,00000	4,43000	5,00000	4,50000	8,60000	4,00000		
26	8,67000	7,33000	5,50000	3,90000	7,00000	8,60000	4,00000		
27	6,00000	8,50000	4,50000	2,67000	3,50000	8,60000	4,00000		
28	6,00000	5,50000	4,00000	10,33000	6,33000	8,60000	4,00000		
29	5,25000	3,60000	3,33000	3,25000	5,00000	8,60000	4,00000		
30	7,00000	3,20000	6,67000	5,67000	5,00000	8,60000	4,00000		
31	6,00000	6,00000	7,20000	3,50000	4,50000	8,60000	4,00000		
32	3,00000	3,60000	5,56000	4,50000	5,71000	8,60000	4,00000		
33	6,50000	7,67000	5,80000	9,00000	6,00000	8,60000	4,00000		
34	3,50000	5,95000	6,75000	3,00000	7,33000	8,60000	4,00000		
35	6,19000	4,00000	4,00000	5,00000	4,50000	8,60000	4,00000		
36	5,17000	4,20000	5,50000	3,90000	7,00000	8,60000	4,00000		
37	8,67000	7,33000	5,50000	3,90000	7,00000	8,60000	4,00000		
38	6,00000	8,50000	4,50000	2,67000	3,50000	8,60000	4,00000		
39	6,00000	5,50000	4,00000	10,33000	6,33000	8,60000	4,00000		
40	5,25000	3,60000	3,33000	3,25000	5,00000	8,60000	4,00000		
41	7,00000	3,20000	6,67000	5,67000	5,00000	8,60000	4,00000		
42	6,00000	6,00000	7,20000	3,50000	4,50000	8,60000	4,00000		

Tabelle: Matrix der Distanz-Koeffizienten. Die 40 Wörter sind alphabetisch geordnet. (WILLKÜRLISTE)

28	7.00000	6.24000	7.40000	6.60000	5.13000	7.67000	7.00000	9.50000	4.50000	12.50000
	6.33000	10.10000	7.00000	3.60000	13.00000	11.50000	7.00000	7.50000	8.80000	6.60000
29	3.85000	11.20000	6.70000	6.50000	5.43000	8.70000	5.40000	10.33000	6.00000	6.33000
	13.30000	6.00000	9.20000	6.75000	3.25000	5.20000	11.00000	6.25000	4.50000	10.50000
	9.00000	8.70000	6.33000	7.00000	10.83000	9.50000	8.90000	6.50000		
30	6.58000	1.00000	6.18000	8.20000	3.60000	4.67000	8.50000	3.00000	3.00000	2.57000
	5.75000	9.50000	4.20000	1.75000	4.13000	6.60000	5.25000	3.20000	6.00000	10.00000
	5.29000	3.00000	6.60000	4.80000	5.40000	4.80000	5.33000	5.25000		
31	8.00000	4.75000	2.20000	8.50000	2.46000	6.00000	5.22000	5.40000	6.20000	
	8.00000	4.75000	2.20000	8.50000	2.46000	6.00000	5.22000	5.40000	6.20000	8.25000
	2.00000	3.67000	5.00000	4.80000	7.00000	2.80000	5.00000	4.20000	5.00000	4.50000
	8.50000	3.70000	3.67000	6.50000	4.16000	3.60000	4.23000	10.25000	6.33000	5.75000
32	10.25000	5.50000	4.67000	5.40000	4.28000	7.00000	4.00000	4.00000	6.60000	4.25000
	6.00000	4.33000	4.67000	6.50000	9.00000	5.30000	2.80000	5.30000	10.33000	4.50000
	4.00000	5.94000	4.25000	5.71000	6.43000	3.60000	5.30000	5.00000	5.00000	5.25000
33	7.00000	4.75000	4.67000	6.50000	6.20000	1.40000	6.00000	6.50000	5.33000	3.67000
	12.50000	5.63000	4.00000	9.67000	7.00000	2.67000	6.00000	3.00000	7.33000	7.00000
	8.00000	10.20000	4.00000	7.50000	2.33000	5.60000	5.00000	11.20000	4.20000	4.00000
34	7.00000	10.0000	4.00000	9.00000	6.00000	3.80000	8.75000	6.04000	10.00000	6.10000
	3.65000	8.70000	10.40000	9.00000	6.00000	5.50000	3.00000	9.80000	3.50000	11.60000
	5.94000	3.00000	5.88000	6.50000	8.33000	5.90000	3.00000	6.04000	3.67000	
	10.15000	8.00000	5.00000	6.35000	6.50000	5.95000	4.00000			
35	4.00000	6.00000	4.00000	4.00000	4.00000	9.00000	5.60000	7.67000	2.00000	10.25000
	6.00000	3.50000	5.20000	5.80000	4.75000	10.70000	5.50000	3.20000	4.50000	8.50000
	9.00000	5.70000	3.00000	8.67000	4.67000	3.00000	5.38000	4.33000	4.33000	6.33000
36	2.25000	9.50000	7.00000	5.86000	6.75000	4.00000	5.27000	6.50000	5.00000	5.46000
	5.60000	3.00000	3.50000	4.67000	5.86000	3.00000	2.25000	5.50000	3.00000	5.34000
	7.00000	5.13000	5.17000	6.00000	7.00000	5.30000	6.00000	4.50000	5.83000	2.50000
37	2.67000	5.45000	5.20000	5.53000	6.50000	6.50000	3.33000	4.50000	3.00000	11.50000
	8.75000	5.81000	2.00000	5.67000	5.75000	5.50000	5.33000	7.25000	6.33000	5.00000
	8.00000	3.50000	6.60000	11.67000	6.25000	9.00000	5.00000	2.50000	10.33000	5.00000
	4.67000	4.00000	11.50000	10.17000	6.00000	5.56000				
38	7.00000	6.00000	7.00000	3.07000	5.33000	6.89000	4.70000	6.00000	3.80000	6.47000
	8.00000	6.09000	9.50000	10.07000	3.50000	6.06000	4.00000	9.33000	7.50000	6.50000
	3.80000	6.33000	7.25000	11.50000	7.67000	9.33000	9.33000	9.33000	4.50000	4.00000
39	1.00000	4.00000	6.34000	6.47000	7.00000	3.90000	6.67000	9.50000	5.00000	7.17000
	3.33000	4.50000	7.20000	5.71000	4.40000	9.25000	3.40000	7.75000	8.00000	8.00000
	3.33000	6.00000	5.57000	6.83000	3.75000	5.00000	6.50000	4.75000	4.00000	6.00000
40	8.80000	5.20000	7.00000	6.16000	5.40000	5.91000	4.00000	7.00000	1.50000	4.80000
	8.33000	10.50000	3.50000	3.60000	4.20000	10.00000	6.25000	3.75000	5.00000	4.67000
	4.50000	3.00000	6.75000	3.40000	1.50000	3.00000	5.80000	5.00000	12.67000	2.00000
	4.50000	4.50000	4.14000	4.40000	3.00000	10.00000	5.00000	4.00000	4.71000	
	1.00000	6.00000	5.33000	5.55000	4.50000	4.92000	4.00000	8.95000		

CLUSTERLISTE

		1	2
1 Beerdigung	1	-.9510	-.7346
2 Begriff	2	-.1931	.5662
3 Buchstabe	3	-.1721	1.0061
4 Busen	4	-.1119	-.5800
5 Dachstuhl	5	-.0561	.2309
6 Durchfall	6	-1.0329	.8536
7 Eisen	7	1.2849	.1601
8 Fuß	8	.5823	-.4917
9 Gabel	9	1.0168	-.2833
10 Geschirr	10	.8939	-.0291
11 Gestank	11	-.4476	-.8947
12 Gott	12	-1.1997	-.1550
13 Hand	13	.8828	-.7074
14 Haus	14	.9416	.3292
15 Hölle	15	-.9554	-.4175
16 Holz	16	1.3526	-.9400
17 Kanne	17	.4041	.0262
18 Kirche	18	-.4199	-1.2797
19 Klopapier	19	-.8746	.4141
20 Kot	20	-.9439	.2884
21 Kuß	21	.0706	-.3703
22 Liebe	22	.0740	-1.1019
23 Material	23	.8611	.7719
24 Orgasmus	24	.3808	-.6556
25 Penis	25	.1237	-.5128
26 Prüfung	26	.4995	.9081
27 Psychologie	27	-.2581	.7726
28 Sarg	28	-.8140	-.8641
29 Satz	29	-.0030	1.2804
30 Schornstein	30	.2711	.1828
31 Sterben	31	-.9109	.6215
32 Stoff	32	.3727	.5665
33 Studium	33	-.4025	.7787
34 Sünde	34	-1.3579	-.1688
35 Teller	35	.9347	-.1221
36 Tod	36	-.7729	-.3654
37 Universität	37	-.3686	.1771
38 Wort	38	.0102	1.1307
39 Zärtlichkeit	39	.0778	-.9276
40 Zimmer	40	1.1309	.5686
	MEAN	0.0000	-.0000
	SIGMA	.7386	.6741

Tabelle: Die Koordinaten der 40 Elemente in der zweidimensionalen Euklidischen MINISSA-Lösung (CLUSTERLISTE).

WILLKÜRLISTE

		1	2
1 Abbildung	1	1.1869	-.3206
2 Abgrund	2	.4004	.5956
3 Amputation	3	-.0466	-.2573
4 ausgeglichen	4	-.9708	-.3920
5 Bordell	5	-.2501	.4598
6 Buchstabe	6	1.3232	-.6351
7 Ferien	7	-.7368	-.3297
8 Folter	8	-.0753	-1.0214
9 Geborgenheit	9	-.7294	.3422
10 Gegend	10	.2810	-1.1629
11 geisteskrank	11	-.7213	.7732
12 Geld	12	.2461	.9260
13 Heizung	13	.6589	-.4060
14 Kante	14	.3941	1.1441
15 Kirche	15	-.5091	-1.2377
16 kotzen	16	.2812	-.8401
17 Krieg	17	-.4007	-.5675
18 Lampe	18	.6288	-.7169
19 Leiche	19	.7612	.6293
20 Luft	20	-1.2974	-.2503
21 Material	21	-1.1473	.2738
22 Musik	22	-.7595	-.5639
23 Orgasmus	23	.4285	.0521
24 Prüfung	24	1.2383	-.0106
25 Psychologie	25	.4793	.4593
26 Rad	26	-.4868	-.7977
27 rechteckig	27	.9831	.5047
28 Regal	28	-.5082	1.3967
29 Regenschirm	29	.6829	1.1024
30 Sau	30	.1653	-.2597
31 schwanger	31	-.1269	.2182
32 sehen	32	-.9146	.0066
33 Sprache	33	.0656	-.8007
34 Stiefel	34	1.1848	.1852
35 Streit	35	-.0058	.8914
36 Tränen	36	.2906	-.3633
37 Treppe	37	-1.0056	.6497
38 Unterschied	38	-.8847	-.9159
39 Verbrechen	39	-.5475	1.0043
40 Wohnung	40	-.3556	.2345
	MEAN	-.0000	.0000
	SIGMA	.7232	.6907

Tabelle: Die Koordinaten der 40 Elemente in der zweidimensionalen Euklidischen MINISSA-Lösung (WILLKÜRLISTE).

Literatur

- AMELANG, M., WENDT, W. & FRÜNDT, H.: Zum Einfluß von Extra-version/Introversion auf Konsolidierungsprozesse beim Behalten verbalen Materials. In: Zeitschrift für experimentelle und angewandte Psychologie, 1977, 24, 525 - 545
- AMSTER, H.: Evaluative judgment and recall in incidental learning. In: J. of Verb. Learn. Verb. Behav., 3, 1964, 466 - 473
- ANISFELD, M.: Learning pairs of nonsense syllables and pleasant and unpleasant adjectives. In: Journal of General Psychology, 1970, 82, 195 - 198
- ANISFELD, M. & LAMBERT, W.E.: When are pleasant words learned faster than unpleasant words. In: J. of Verb. Learn. Verb. Behav., 5, 1966, 132 - 141
- AUGUST, G.J.: Affective meaningfulness, self-concept and the verbal learning styles of white and black children. In: Dissertation Abstracts International, 1976, 36 (9-B), 4717
- AUGUST, G.J. & FELKER, D.W.: Role of affective meaningfulness and self-concept in the verbal learning styles of white and black children. In: Journal of Educational Psychology, 1977, 69, 253 - 260
- AUGUST, G.J., RYCHLAK, J.F. & FELKER, D.W.: Affective assessment, self-concept, and the verbal learning styles of fifth-grade children. In: Journal of Educational Psychology, 1975, 67, 801 - 806
- BADDELEY, A.D.: Die Psychologie des Gedächtnisses. Stuttgart: Klett-Cotta, 1979
- BASSOS, C.A.: Affective content and contextual constraint in recall by paranoid, and nonpsychiatric patients. In: Journal of consulting and Clinical Psychology, 1973, 40, 126 - 132
- BECKER-CARUS, C. HEYDEN, T. & ZIEGLER, G.: Psychophysiologische Methoden. Stuttgart: Ferdinand Enke Verlag, 1979
- BERLYNE, D.E., BORSA, D.M., CRAW, M.A., GELMAN, R.S. & MANDEL, E.E.: Effects of stimulus complexity and induced arousal on paired-associate learning. In: J. of Verb. Learn. Verb. Behav., 4, 1965, 291 - 299
- BERLYNE, D.E. & CAREY, S.T.: Incidental learning and the timing of arousal. In: Psychonomic Science, 1968, 13, 103 - 104
- BERRY, R.N.: Skin conductance levels and verbal recall. In: J. of Experimental Psychology, 1962, 63, 275 - 277

- BOCK, M.: Angenehme und unangenehme Erfahrungen aus gedächtnispsychologischer Sicht. Antrittsvorlesung am 28.6.1978 in der Abteilung für Philosophie, Pädagogik, Psychologie der Ruhr-Universität Bochum, 1978 (a)
- BOCK, M.: Wort-, Satz- und Textverarbeitung. Stuttgart: Kohlhammer 1978 (b)
- BOTTENBERG, E.H.: Emotionspsychologie. München, 1972
- BROWN, C.H.: The relation of magnitude of galvanic skin responses and resistance levels to the rate of learning. In: J. of Experimental Psychology, 1937, 20, 262 - 278
- BROWN, W.P.: A retrospective study of stimulus variables in word association. In: J. of Verb. Learn. Verb. Behav., 10, 1971, 355 - 366
- BROWN, W.P. & URE, D.M.: Five rated characteristics of 650 word association stimuli. In: Br. J. Psychology, 1969, 60, 233 - 249
- BUTTER, M.J.: Differential recall of paired-associates as a function of arousal and concreteness-imagery levels. In: J. of Experimental Psychology, 1970, 84, 252 - 256
- CAMP, B.W. & ZIMET, S.G.: Recall of aggressive words by kindergarten children. In: Psychological Reports, 1977, 33, 575 - 578
- CATTELL, R.B.: Die empirische Erforschung der Persönlichkeit. Weinheim: Beltz, 1973
- CLAUSS, G. (Hrsg.): Wörterbuch der Psychologie. Köln: Pahl-Rugenstein, 1976
- CLAUSS, G. & EBNER, H.: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen. Frankfurt a.M. und Zürich: Harri Deutsch, 1970
- COFER, C.N.: Properties of verbal materials and verbal learning. In: KLING, J.W. & RIGGS, L.A. (Hrsg.): Experimental Psychology. Holt, Rinehart & Winston, 1971, 847 - 904
- CORTEEN, R.S.: Skin conductance changes and word recall. In: Br. J. of Psychology, 1969, 60, 81 - 84
- CRAMER, P.: Word association. New York: Academic Press, 1968
- DUKES, W.F. & BASTIAN, J.: Recall of abstract and concrete words equated for meaningfulness. In: J. of Verb. Learn. Verb. Behav., 1966, 5, 455 - 458
- DUTTA, S. & KANUNGO, R.N.: Affect and memory: A reformulation. Oxford: Pergamon Press, 1975

- EDELBERG, R.: Electrical activity of the skin. In: GREENFIELD, N.S. & STERNBACH, R.A. (Hrsg.): Handbook of Psychophysiology. New York: Holt, Rinehart & Winston, 1972, 367 - 418
- EHRHARDT, K.J.: Neuropsychologie 'motivierten' Verhaltens. Stuttgart: Enke, 1975
- EISDORFER, C.: Arousal and Performance: Experiments in verbal learning and a tentative theory. In: TALLAND, G.A. (Hrsg.): Human aging and behavior. New York: Academic Press, 1968, 189 - 216
- EMRICH, H. & HEINEMANN, L.G.: EEG bei unterschwelliger Wahrnehmung emotional bedeutsamer Wörter. In: Psychologische Forschung, 1966, 29, 285 - 296
- EYSENCK, H.J.: The measurement of emotion: Psychological Parameters and methods. In: LEVI, L. (Hrsg.): Emotions - their parameters and Measurement. New York: Raven Press Publ., 1975, 439 - 467
- EYSENCK, M.W.: Effects of noise, activation level, and response dominance on retrieval from semantic memory. In: J. of Experimental Psychology: Human Learning and Memory, 1975, 104, 143 - 148
- EYSENCK, M.W.: Arousal, learning, and memory. In: Psychological Bulletin, 1976, 83, 389 - 404
- EYSENCK, M.W.: Human memory. Theory, research and individual differences. Oxford: Pergamon Press, 1977
- FRINCKE, G.: Word characteristics, associative-relatedness, and the free-recall of nouns. In: J. of Verb. Learn. Verb. Behav., 1968, 7, 366 - 372
- FRÖBES, J.: Lehrbuch der experimentellen Psychologie. 2. Bd. Freiburg im Breisgau: Herder, 1929
- FUCHS, A.: Grundzüge einer Verhaltenstheorie der Bedeutung. In: BERGLER, R.: Das Eindrucksdifferential. Bern: Huber, 1975 (a), 33 - 68
- FUCHS, A.: Das Eindrucksdifferential als Instrument zur Erfassung emotionaler Bedeutungsprozesse. In: BERGLER, R. (s.o.), 1975 (b), 69 - 100
- GORDON, R.A.: An empirical study of aggressive word recall of severely disturbed primary aged children. In: Dissertation Abstracts International, 1976, 36 (12-A), 7947
- GREGG, V.: Word frequency, recognition and recall. In: BROWN, J. (Hrsg.): Recall and recognition. London, New York, 1976, 183 - 216
- GRUNEBERG, M.M.: The serial position curve and the distinction between short- and long-term memory. In: Acta Psychologica, 1972, 36, 221 - 225

- GUERNETTE, J.G.: The relation between rate of learning and verbal, affective stimuli. In: Dissertation Abstracts International, 1975, 35 (8-A), 5116
- HAGGARD, E.A.: On the application of analysis of variance to GSR data: Some effects of the use of inappropriate measures. In: J. of Experimental Psychology, 1949, 39, 861 - 867
- HALL, J.F.: Verbal learning and retention. Philadelphia, New York, Toronto: J.B.Lippincott C., 1971
- HEINRICH, B.A.: Motivation and long-term memory. In: Psychonomic Science, 1968, 12 (4), 149 - 150
- HEYDEN, T.: Elektrodermale Aktivität (EDA). In: BECKER-CARUS, C., HEYDEN, T. & ZIEGLER, G.: Psychophysiologische Methoden. Stuttgart: Ferdinand Enke Verlag, 1979, 83 - 97
- HOFSTÄTTER, R.R. (Hrsg.): Psychologie. Fischer Lexikon. Frankfurt a.M.: Fischer, 1972
- HOFSTÄTTER, P. & WENDT, D.: Quantitative Methoden der Psychologie. Bd.1, Frankfurt a.M.: Barth, 1974
- HOHN, R.L., JOHNSEN, E.P. & TRACY, D.B.: Arousal and imagery effects on reminiscence and forgetting of nouns. In: J. of General Psychology, 1976, 94, 147-148
- HOLMES, D.S.: Repression or interference? A further investigation. In: J. of Personality and Social Psychology, 1972, 22, 163 - 170
- HOLMES, D.S.: Investigation of repression: differential recall of material experimentally or naturally associated with Ego threat. In: Psychological Bulletin, 1974, 81, 632 - 653
- HÖRMANN, H.: Psychologie der Sprache. Berlin, Heidelberg: Springer, 1977
- HÖRMANN, H. & OSTERKAMP, U.: Über den Einfluß von kontinuierlichem Lärm auf die Organisation von Gedächtnisinhalten. In: Zeitschrift für angewandte Psychologie, 1966, 13, 31 - 38
- HOWARTH, E. & EYSENCK, H.J.: Extraversion, arousal, and paired-associate recall. In: J. of Experimental Research in Personality, 1968, 3, 114 - 116
- HOWE, E.S.: Number of different free associates: A general measure of associative meaningfulness. In: J. of Verb. Learn. Verb. Behav., 1972, 11, 18-28
- HURLOCK, E.B.: Die Entwicklung des Kindes. Weinheim: Beltz 1970
- HUTT, L.D. & ANDERSON, J.P.: The relationship between pupil size and recognition threshold. In: Psychonomic Science, 1967, 9, 477 - 478

- JEFFMAR, M.: The effects of attention and repetition on the organization of recall of randomly arranged associates. In: J. of General Psychology, 1975, 92, 245 - 253
- JONES, H.E.: The galvanic skin reflex in Infancy. In: Child Development, 1930, 1, 106 - 110
- JONES, H.E. & WECHSLER, D.: Galvanometric technique in studies of association. In: American J. of Psychology, 1928, 40, 607 - 612
- KAEDING, F.W.: Häufigkeitwörterbuch der deutschen Sprache. 1898
- KAPLAN, R. & KAPLAN, S.: The arousal-retention interval interaction revisited: The effects of some procedural changes. In: Psychonomic Science, 1969, 15, 84 - 85
- KAPLAN, S. & KAPLAN, R.: Arousal and memory: A comment. In: Psychonomic Science, 1968, 10, 291 - 292
- KAPLAN, S. & KAPLAN R.: The interaction of arousal and retention interval: Ipsative vs. normative scoring. In: Psychonomic Science, 1970, 19, 115 - 116
- KAPLAN, S., KAPLAN R. & SAMPSON, J.R.: Encoding and arousal factors in free recall of verbal and visual material. In: Psychonomic Science, 1968, 12, 73 - 74
- KAYTON, L. & KOH, S.D.: Hypohedonia in schizophrenia. In: J. of Nervous and Mental Disease, 1975, 161 (6), 412 - 420
- KELLEY, R.J.: The learning and recall of emotional and neutral words in a negative transfer situation as a function of differentiation. In: Dissertation Abstracts, 1968, 28 (11-B), 4757-4758
- KERLINGER, F.N.: Foundations of behavioral research. London: Holt, Rinehart & Winston, 1970
- KINTSCH, W.: Memory and cognition. New York: John Wiley & Sons, 1977
- KLEINSMITH, L.J. & KAPLAN, S.: Paired-associate learning as a function of arousal and interpolated interval. In: J. of Experimental Psychology, 1963, 65, 190 - 193
- KLEINSMITH, L.J. & KAPLAN, S.: Interaction of arousal and recall interval in nonsense syllable paired-associate learning. In: J. of Experimental Psychology, 1964, 67, 124 - 126
- KLEINSMITH, L.J., KAPLAN, S. & TARTE, R.D.: The relationship of arousal to short- and long-term verbal recall. In: Canadian J. Psychology, 1963, 17, 393 - 397

- KLING, J.W. & RIGGS, L.A. (Hrsg.): Experimental Psychology. Holt, Rinehart & Winston, 1971
- KUSA, D.: Effect of word meaningfulness and frequency on short-term recognition. In: *Studia Psychologica*, 1976, 18, 197 - 206
- LANC, O.: Psychophysiologische Methoden. Stuttgart: Kohlhammer, 1977
- LEICHT, K.L.: Recall and judged frequency of implicitly occurring words. In: *J. of Verb. Learn. Verb. Behav.*, 1968, 7, 918 - 923
- LEVI, L. (Hrsg.): Emotions. Their parameters and measurement. New York: Raven Press Publ., 1975
- LEVONIAN, E.: Attention and consolidation as factors in retention. In: *Psychonomic Science*, 1966, 6, 275 - 276
- LEVONIAN, E.: Retention over time in relation to arousal during learning: an explanation of discrepant results. In: *Acta Psychologica*, 1972, 36, 290 - 321
- LLOYD, K.E.: Short-term retention as a function of word frequency. In: *J. of Verb. Learn. Verb. Behav.* 1964, 3, 286 - 289
- LUSTIG, R.S. & WITRYOL, S.L.: The influence of differential incentive motivation on elaboration of semantic processing in free recall. In: *J. of General Psychology*, 1977, 96, 321 - 322
- MALTZMAN, I., KANTOR, W. & LANGDON, B.: Immediate and delayed retention, arousal, and the orienting and defensive reflexes. In: *Psychonomic Science*, 1966, 6, 445 - 446
- MANNING, S.K. & GOLDSTEIN, F.D.: Recall of emotional and neutral words as a function of rate and organization of list presentation. In: *J. of General Psychology*, 1976, 95, 242 - 249
- MANNING, S.K. & JULIAN, L.: Recall of emotional words. In: *J. of General Psychology*, 1975, 92, 237 - 244
- MANNING, S.K. & MELCHIORI, M.P.: Words that upset urban college students: Measurement with GSR, and rating scales. In: *J. of Social Psychology*, 1974, 94, 305 - 306
- MARX, W.: Statistische Information und Assoziative Bedeutung verschiedener Wortarten. In: *Zeitschrift für experimentelle und angewandte Psychologie*, 1978, 25, 431 - 440
- MARX, W.: Eine Methode zur geometrischen Darstellung des Clusterings bei Free-Recall-Daten. In: *Arch. Psychologie*, 1980, 132, 165 - 173

- Mc CURDY, H.G.: Consciousness and the galvanometer. In: Psychological Review, 1950, 57, 322 - 327
- McGEOCH, J.A. & IRION, A.L.: The psychology of human learning. New York: Longmans, Green & Co., 1952
- McLEAN, P.D.: Induced arousal and time of recall as determinants of paired-associate recall: In: British J. Psychology, 1969, 60, 57 - 62
- McNULTY, J.A. & ISNOR, C.J.: Affective connotation of words a factor in free recall learning. In: British J. Psychology, 1967, 58, 51 - 56
- MEIER, H.: Deutsche Sprachstatistik. Band 2. Hildesheim: Georg Olms Verlagsbuchhandlung, 1967
- MENZIES, R.: The comparative memory values of pleasant, unpleasant and indifferent experiences. In: J. Experimental Psychology, 1935, 18, 267 - 279
- MILNER, P.M.: Physiological Psychology. London: Holt, Rinehart & Winston; 1970
- MOSCU, J.: Recherches psychophysiologiques sur l'émotivité chez les sportifs d'élite. In: International J. of Sport Psychology, 1971, 2, 109 - 117
- MOSCU, J.: Experimental investigations on the emotivity of top sportsmen. In: Revue Roumaine des Sciences Sociales, Serie de Psychologie, 1972, 16, 29 - 40
- MOSCU, J.: L'étude expérimentale du trac chez les acteurs. In: Revue Roumaine des Sciences Sociales, Séries de Psychologie, 1973. 17, 145 - 158
- MURRAY, F.S.: Effects of narrative stories on recall. In: Bulletin of the Psychonomic Society, 1974, 4 (6), 577 - 579
- NOBLE, C.E.: Emotionality (e) and meaningfulness (m). In: Psychological Reports, 1958, 4, 16
- NOBLE, C.E.: Measurement of association value (a), rated associations (a') and scaled meaningfulness (m') for the 2100 CVC combinations of the English Alphabet. In: Psychological Reports, 1961, 8, 487 - 521
- OSGOOD, C.E.: Method and theory in experimental Psychology. New York: University Press, 1964
- OSGOOD, C.E., SUCI, G.J. & TANNENBAUM, P.H.: The measurement of meaning. Urbana, Chicago, London: University of Illinois Press, 1957/1967
- OTT, T. & MATTHIES, H.: Lernen und Gedächtnis. In: Psychologie des 20. Jahrhunderts, Bd.6, Darmstadt: Kindler, 1978, 988 - 1018

- PAIVIO, A.: Abstractness, imagery, and meaningfulness in paired-associate learning. In: J. of Verb. Learn. Verb. Behav., 1965, 4, 32 - 38
- PAIVIO, A.: A factor-analytic study of word attributes and verbal learning. In: J. of Verb. Learn. Verb. Behav., 1968, 7, 41 - 49
- PAIVIO, A.: Mental imagery in associative learning and memory. In: Psychological Review, 1969, 76, 241 - 263
- PAIVIO, A. & YUILLE, J.: Mediation instructions and word attributes in paired-associate learning. In: Psychonomic Science, 1967, 8, 65 - 66
- PAIVIO, A. YUILLE, J.C. & MADIGAN, S.A.: Concreteness, imagery, and meaningfulness values for 925 nouns. In: J. of Experimental Psychology, Monograph Supplement, 1968, 76, 1 - 9
- PAIVIO, A. YUILLE, J.C. & ROGERS, T.B.: Noun imagery and meaningfulness in free recall and serial recall. In: J. of Experimental Psychology, 1969, 79, 509 - 514
- PALLAK, M.S. & ANDREW, J.: The effects of expected shock and expected payment on incidental retention. In: Psychonomic Science, 1970, 18, 323 - 325
- PETERSON, F. & JUNG, C.G.: Psycho-physical investigations with the galvanometer and pneumograph in normal and insane individuals. In: Brain, 1907, 30, 153 - 218
- PETERSON, L.R.: Verbal learning and memory. In: Ann. Rev. Psychology, 1977, 28, 393 - 415
- POWERS, M.: Scaling words on degree of arousal, and short- and long-term retention. In: Dissertation Abstracts International, 1974, 34 (8-A, PT 1), 4884
- RAPAPORT, D.: Gefühl und Erinnerung. Stuttgart: Ernst Klett, 1977 (Übersetzung von 'Emotions and memory' 1971, 2. Auflage 1950)
- REYKOWSKI, J.: Psychologie der Emotionen. Donauwörth: Ludwig Auer, 1973
- RHONE, T.A.: Learning, recall, and relearning of verbal, affective stimuli. In: Dissertation Abstracts International, 1975, 35 (9-A), 5937
- ROUTH, D.A. & ELLIS, S.: The processing of affectively-toned words in incidental learning: A test of differential acquisition and response bias hypotheses. In: Acta Psychologica, 1975, 39, 51 - 67
- RUNDLE, S.A.: An experimental and theoretical study of recognition memory. In: Dissertation Abstracts International, 1970, 30 (11-B), 5265

- RYCHLAK, J.F., CARLSEN, N.L. & DUNNING, L.P.: Personal adjustment and the free recall of materials with affectively positive and negative meaningfulness. In: J. of Abnormal Psychology, 1974, 83, 480 - 487
- SAMPSON, J.R.: Further study of encoding and arousal factors in free recall of verbal and visual material. In: Psychonomic Science, 1969, 16, 221 - 222
- SCHMITT, J.C. & FORRESTER, W.E.: Effects of stimulus concreteness-imagery and arousal on immediate and delayed recall. In: Bulletin of the Psychonomic Society, 1973, 2, 25 - 26
- SCHÖNPFLUG, W.: Über Aktivationsprozesse im Lernversuch. Frankfurt a.M.: Kramer, 1963
- SCHÖNPFLUG, W.: Paarlernen, Behaltensdauer und Aktivierung. In: Psychologische Forschung, 1966, 29, 132-148
- SCHÖNPFLUG, W. (Hrsg.): Methoden der Aktivierungsforschung. Bern: Hans Huber, 1969
- SCHÖNPFLUG, W. & BEIKE, P.: Einprägen und Aktivierung bei gleichzeitiger Variation der Absichtlichkeit des Lernens und der Ich-Bezogenheit des Lernstoffs. In: Psychologische Forschung, 1964, 27, 366 - 376
- SCHULTER, G.: Einprägungsaktivität und Gedächtnisleistung. In: Archiv für Psychologie, 1975, 127, 217-283
- SCHUTZ, S.R. & KEISLAR, E.R.: Young children's immediate memory of word classes in relation to social class. In: J. of Verb. Learn. Verb. Behav., 1972, 11, 13 - 17
- SEAMON, J.G. & MURRAY, P.: Depth of processing in recall and recognition memory: Differential effects of stimulus meaningfulness and serial position. In: J. of Experimental Psychology: Human Learning and Memory, 1976, 2, 680 -687
- SIEGEL, S.: Nichtparametrische statistische Methoden. Frankfurt a.M.: Fachbuchhandlung für Psychologie, 1976
- SILVERSTEIN, A. & DIENSTBIER, R.A.: Rated pleasantness and association value of 101 English nouns. In: J. of Verb. Learn. Verb. Behav., 1968, 7, 81-86
- SINHA, M. & SINHA, L.N.: A study of learning of emotional and non-emotional words. In: Indian Psychological Review, 1977, 15, 43 - 45
- SMITH, E.E. & SLOBODA, W.: Some relations between affect and recall: An exploratory study. In: Psychonomic Science, 1968, 12, 257 - 258

- SMITH, M.G. & HARLESTON, B.W.: Stimulus abstractness and emotionality as determinants of behavioral and physiological responses in a word-association task. In J. of Verb. Learn. Verb. Behav., 1966, 5, 309 - 313
- SMITH, W.: Experiment on memory and affective tone. In: British Journal of Psychology, 1921, 11, 236 - 250
- SMITH, W.: A note on the use of the PGR. In: The British J. of Psychology, 12, 1921, 282 - 287
- SMITH, W.: The measurement of emotion. Maryland: McGrath 1970, Neuauflage, (London 1922)
Die Neuauflage wurde fälschlicherweise unter dem Namen CARINGTON, W. herausgebracht.
- SMYTHE, P.C. & PAIVIO, A.: A comparison of the effectiveness of word imagery and meaningfulness in paired-associated learning of nouns. In: Psychonomic Science, 1968, 10, 49 - 50
- SPREEN, O. & SCHULZ, R.W.: Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. In: J. of Verb. Learn. Verb. Behav., 1966, 5, 459 - 468
- SYZ, H.C.: Psycho-galvanic studies on sixty-four medical students. In: British Journal of Psychology, 1926/27, 17, 54 - 69 (a)
- SYZ, H.C.: Observations on the unreliability of subjective reports of emotional reactions. In: British J. of Psychology, 1926/27, 17, 119 - 126 (b)
- THAYER, R.E.: Measurement of activation through self-report. In: Psychological Reports, 1967, 20, 663 - 678
- TEASDALE, J.D. & FOGARTY, S.J.: Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. In: J. of Abnormal Psychology, 1979, 88, 248 - 257
- TEMPLER, D.I.: The relationship between verbalized and nonverbalized death anxiety. In: J. of Genetic Psychology, 1971, 119, 211 - 214
- THOMSON, R.H.: An experimental study of memory as influenced by feeling tone. In: J. of Experimental Psychology, 1930, 13, 462 - 468
- TRAXEL, W.: Die Möglichkeit einer objektiven Messung der Stärke von Gefühlen. In: Psychologische Forschung, 1960, 26, 75 - 90
- TRAXEL, W.: Gefühl und Gefühlsausdruck. In: MEILI, R. & ROHRACHER, H.: Lehrbuch der experimentellen Psychologie. Bern, Stuttgart: Hans Huber, 1963, 215 - 263

- TRAXEL, W. & BECHER, S. : Beiträge zur Meßmethodik und Interpretation von Hautwiederstandsänderungen. In: Zeitschrift für Psychologie, 1957, 160, 282 - 301
- TURNER, R.H. & BARLOW, J.A.: Memory for pleasant and unpleasant experiences: some methodological considerations. In: J. of Experimental Psychology, 1951, 42, 189 - 196
- UEHLING, B.S.: Arousal in verbal learning. In: DUNCAN, C.P., SECHREST, L. & MELTON, A.W. (Hrsg.): Human memory: Festschrift for Benton J. Underwood. New York: Meredith Corporation, 1972, 263-282
- UEHLING, B.S. & SPRINKLE, R.: Recall of a serial list as a function of arousal and retention interval. In: J. of Experimental Psychology, 1968, 78, 103 - 106
- VENABLES, P.H. & MARTIN, I.: Skin resistance and skin potential. In: VENABLES, P.H. & MARTIN, I. (Hrsg.): A manual of psychophysiological Methods. Amsterdam: North-Holland Publishing C., 1967, 51-102
- VICORY, A.C. & ASHER, J.J.: A simplified associative-frequency measure of meaningfulness in verbal learning. In: J. of Verb. Learn. Verb. Behav., 1966, 5, 507 - 513
- WALKER, E.L.: Action decrement and its relation to learning. In: Psychological Review, 1958, 65, 129 - 142
- WALKER, E.L. & TARTE, R.D.: Memory storage as a function of arousal and time with homogeneous and heterogeneous lists. In: J. of Verb. Learn. Verb. Behav., 1963, 2, 113 - 119
- WALSCHBURGER, P.: Zur Standardisierung und Interpretation elektrodermalen Meßwerte in psychologischen Experimenten. In: Zeitschrift für experimentelle und angewandte Psychologie, 1975, 22, 514 - 533
- WALSCHBURGER, P.: Zur Beschreibung von Aktivierungsprozessen. Dissertation, Freiburg i.Br., 1976
- WARREN, L.R. & HARRIS, L.J.: Arousal and memory: Phasic Measures of arousal in a free recall task. In: Acta Psychologica, 1975, 39, 303 - 310
- WECHSLER, D.: The measurement of emotional reactions. Researches of the galvanic reflex. In: Archives of Psychology, 1925, No.76, 1 - 181
- WECHSLER, D. & JONES, H.E.: A study of emotional specificity. In: American J. of Psychology, 1928, 40, 600 - 606

- WESNER, C.E.: Induced arousal and word-recognition learning by mongoloids and normals. In: Perceptual and Motor Skills, 1972, 35, 586
- WOODWORTH, R.S. & SCHLOSBERG, H.: Experimental Psychology. New York: Holt, Rinehart & Winston, 1965
- YARMEY, A.D.: Word arousal and verbal mediation. In: Psychonomic Science, 1966, 6, 451 - 452
- ZELLER, A.F.: An experimental analogue of repression. I. Historical summary. In: The Psychological Bulletin, 1950, 47, 39 - 51 (a)
- ZELLER, A.F.: An experimental analogue of repression. II. The effect of individual failure and success on memory measured by relearning. In: J. of Experimental Psychology, 1950, 40, 411-422 (b)
- ZIPF, G.K.: The Psycho-Biology of language. An introduction to dynamic philology. 1965 (1935)
- ZÖHRER, I.: Über den Einfluß von unlustbetonten Affekterlebnissen auf Wahrnehmung, Lernen und Denken. In: Zeitschrift für experimentelle und angewandte Psychologie, 1970, 17. 670 - 686

LEBENS LAUF

Maria Elisabeth Schürer-Necker, geb. Schürer

16.12.1950	geboren in München als Tochter des Fotografen Hans Schürer und seiner Frau Felizitas, geb. Spurk
1957 - 1961	Volksschule der Englischen Fräulein in München
1961 - 1970	Louise-Schröder-Gymnasium in München
Juni 1970	Abitur
WS 1970/71	Studium der Philosophie und Zeitungswissenschaft an der Universität München
SS 71 - WS 76/77	Studium der Psychologie an der Universität München
Dezember 1976	Diplom in Psychologie
SS 77 - WS 78/79	Promotionsstudium Hauptfach: Psychologie Nebenfächer: Soziologie und Psycholinguistik an der Universität München
seit 1.4.1977	Verwalterin einer wissenschaftlichen Assistentenstelle am Institut für Psychologie - Allgemeine Psychologie - der Universität München

